2016

Insights Into Terminal Erythropoiesis Influenced By Human Genetic Variation

Elizabeth Traxler
University of Pennsylvania, liztraxler@gmail.com

Insights Into Terminal Erythropoiesis Influenced By Human Genetic Variation

Abstract
Red blood cells (RBCs) carry hemoglobin, enabling delivery of oxygen to all tissues of the body. They are the products of a highly specialized differentiation process that begins with a hematopoietic stem cell and results in an enucleated, biconcave RBC. This thesis is focused on the use of human genetic studies to gain a better understanding of the molecular processes occurring during terminal erythroid differentiation. We studied the regulation and roles of two erythroid-restricted genes, Trim58 and Hemoglobin Gamma Chain (HBG1 and HBG2, γ-globin), by using a combination of loss-of-function techniques, including RNA-interference-mediated gene suppression, a mutant mouse model, and CRISPR/Cas9 mediated genome editing. Previous genome-wide association studies implicated variation in TRIM58 in RBC development and function. Our experiments with Trim58 revealed a direct interaction with the molecular motor dynein and enzymatic function as an E3 ubiquitin ligase in promoting its proteasomal degradation. This interaction is necessary for enucleation in knockdown studies in vitro, but genetic studies in the mouse show Trim58 is not required for erythropoiesis or enucleation. In the second part of this thesis, we used CRISPR/Cas9 to recreate a known mutation associated with hereditary persistence of fetal hemoglobin, a benign condition that ameliorates co-inherited sickle cell disease (SCD). Genome editing in human hematopoietic stem and progenitor cells reversed the hemoglobin switch at levels sufficient in vitro to correct pathological morphologies in SCD patient-derived RBCs. We identify a cis-regulatory element in the γ-globin promoter as a potential target for genome-editing therapy for SCD. Together, these findings underscore the importance of utilizing both common and rare genetic variants to uncover new aspects of erythroid biology.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Cell & Molecular Biology

First Advisor
Mitchell J. Weiss

Subject Categories
Cell Biology | Genetics | Molecular Biology

INSIGHTS INTO TERMINAL ERYTHROPOIESIS INFLUENCED BY HUMAN GENETIC VARIATION

Elizabeth A. Traxler

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

_____

Dr. Mitchell J. Weiss, MD, PhD, Professor of Pediatrics

Graduate Group Chairperson

_____

Dr. Daniel S. Kessler, PhD, Associate Professor of Cell and Developmental Biology

Dissertation Committee

Mark L. Kahn, MD, Professor of Medicine (Chair)

Erika L. Holzbaur, PhD, Professor of Physiology

Stephen A. Liebhaber, MD, Professor of Genetics

Michael Lampson, PhD, Associate Professor of Biology

Dedication

I dedicate this work to

my father Curt, mother MaryAnn, brothers Shoji, Nikolas, and Alex, and sisters Emily and Fumie for all their love and support;

my partner through it all, Michael;

and my advisor Mitch Weiss for his mentorship and scientific influence.

ACKNOWLEDGMENTS

I would like to thank the following for scientific support:

Weiss Lab members

Current	*Former*
Mitch Weiss	Chris Thom
Yu Yao	Jiyoon Noh
Vikram Paralkar	Jenna Nickas
Christophe Lechauve	Olivia Zhou
Ruopeng Feng	Guowei Zhao
Jean-Yves Metais	Eugene Khandros
Peng Xu	Orna Steinberg-Shemer
Philip Doerfler	Chiaka Aribeana
Kaitly Woodard	

St. Jude Children's Research Hospital

Chunliang Li	Michael Wang
Lance Palmer	Evadnie Rampersaud

Blobel Lab
Gerd Blobel
Sarah Hsu
Jeremy Grevet
Peng Huang

Holzbaur Lab	Kahn Lab
Jacob Lazarus	Patricia Mericko-Ishizuka
Mariko Tokito	

Thesis Committee
Mark L Kahn
Erika L Holzbaur
Michael A Lampson
Steve A Liebhaber

This work was funded by National Institute of Health Grants R01 DK092318 (MJW) and R01 DK61692 (MJW). I was supported by 5F30DK102291-02.

ABSTRACT

INSIGHTS INTO TERMINAL ERYTHROPOIESIS INFLUENCED BY

HUMAN GENETIC VARIATION

Elizabeth A. Traxler

Dr. Mitchell Weiss

Red blood cells (RBCs) carry hemoglobin, enabling delivery of oxygen to all tissues of the body. They are the products of a highly specialized differentiation process that begins with a hematopoietic stem cell and results in an enucleated, biconcave RBC. This thesis is focused on the use of human genetic studies to gain a better understanding of the molecular processes occurring during terminal erythroid differentiation. We studied the regulation and roles of two erythroid-restricted genes, *Trim58* and *Hemoglobin Gamma Chain* (*HBG1* and *HBG2,* γ-globin), by using a combination of loss-of-function techniques, including RNA-interference-mediated gene suppression, a mutant mouse model, and CRISPR/Cas9 mediated genome editing. Previous genome-wide association studies implicated variation in *TRIM58* in RBC development and function. Our experiments with Trim58 revealed a direct interaction with the molecular motor dynein and enzymatic function as an E3 ubiquitin ligase in promoting its proteasomal degradation. This interaction is necessary for enucleation in knockdown studies *in vitro*,

but genetic studies in the mouse show Trim58 is not required for erythropoiesis or enucleation. In the second part of this thesis, we used CRISPR/Cas9 to recreate a known mutation associated with hereditary persistence of fetal hemoglobin, a benign condition that ameliorates co-inherited sickle cell disease (SCD). Genome editing in human hematopoietic stem and progenitor cells reversed the hemoglobin switch at levels sufficient *in vitro* to correct pathological morphologies in SCD patient-derived RBCs. We identify a cis-regulatory element in the γ-globin promoter as a potential target for genome-editing therapy for SCD. Together, these findings underscore the importance of utilizing both common and rare genetic variants to uncover new aspects of erythroid biology.

TABLE OF CONTENTS

List of Tables

List of Illustrations

Preface

This thesis is a series of chapters based on published and submitted manuscripts, and the results are organized and presented closely to how they were published. Each chapter begins with a statement explaining the motivating factors driving the studies, and closes with a brief reflective commentary. Chapter 6 contains more in-depth analysis on the work through a retrospective lens.

Chapter 1 Introduction

1.1 Red blood cell function

Circulating red blood cells (RBC), or erythrocytes, deliver oxygen (O_2) to all tissues of the body. Each mature cell is comprised of ~340 mg/mL hemoglobin (Hb), a heme-containing metalloprotein complex accounting for over 90% of the cellular protein composition (Roux-Dalvai et al., 2008). Hb is a heterotetrameric complex of two α- and two β-like globin subunits or chains, which are each coordinated to a heme. Each heme coordinates an O_2 molecule and allows efficient off-loading to deoxygenated tissues. While Hb is the universal O_2 carrier, packaging high levels of Hb in the form of cells is advantageous for mammalian life in contrast to invertebrate O_2 transport, which relies on via free, high-molecular-weight Hb. If vertebrates also lacked an efficient Hb packaging strategy, extremely high oncotic pressure would preclude a closed circulatory system.

1.2 Erythropoiesis

Erythropoiesis is developmentally divided into three waves, based on erythrocyte size, tissue of origin, and type of hemoglobin expressed. Two main functional requirements highlight the differences between embryonic and adult erythropoiesis. In the embryo, RBCs are required before long-term hematopoietic stem cell (HSC) populations and their environmental niches are established. Additionally, the rapid growth during embryogenesis demands dramatic red cell output (McGrath and Palis,

2008). In contrast, adult erythropoiesis normally remains at steady state in the absence of a hematopoietic stressor such as hemorrhage or hemolysis. The first wave, the primitive wave, begins in the yolk sac and yields RBCs expressing embryonic globin genes (Palis et al., 1995; Silver and Palis, 1997). Until recently, these RBCs were assumed to remain nucleated, but both human and murine primitive erythroblasts enucleate in circulation (McGrath et al., 2008). The second wave consists of erythro-myeloid progenitors from the yolk sac, which can produce RBCs, platelets, and some myeloid lineages (Palis et al., 1999; Wong et al., 1986). These progenitors ultimately migrate and seed the fetal liver mid-gestation for the third wave, definitive erythropoiesis (Lux et al., 2008; McGrath et al., 2015; 2011). Definitive erythropoiesis ultimately takes place in the bone marrow through adulthood, and cells are derived from HSCs. Human definitive RBCs initially express fetal hemoglobin (HbF, $\alpha 2\gamma 2$) and post-natally express adult hemoglobin (HbA, $\alpha 2\beta 2$).

1.3 Terminal erythroid differentiation

Definitive erythroid cells are derived from HSCs, which proliferate and differentiate into all blood lineages. The committed erythroid progenitors are burst-forming unit erythroid (BFU-E) and colony-forming unit erythroid (CFU-E) cells, based on their capacity to form colonies in semi-solid media. CFU-E committed progenitors then undergo 4-5 specialized cell divisions, wherein gene expression profiles and cell morphology change dramatically.

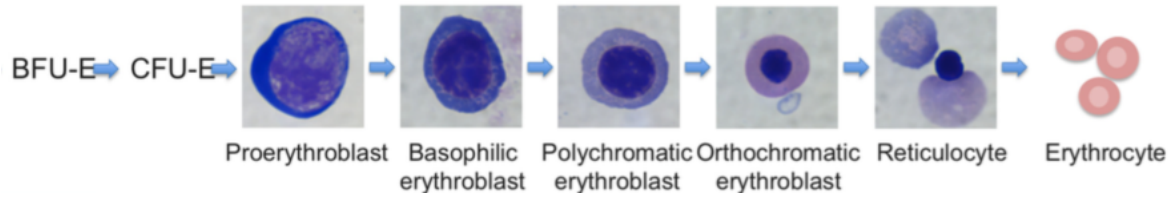


Figure 1.1. Terminal erythropoiesis.

Morphological features of the characteristic progression from erythroid progenitors (BFU-E and CFU-E) to the enucleated reticulocyte and erythrocyte. Micrographs courtesy of James Palis, adapted from Max Pimkin.

Gene expression changes

Several groups have performed global transcriptome analysis at various stages of erythroid differentiation in primary cells (An et al., 2014; Liang et al., 2015; Merryweather-Clarke et al., 2011; Pishesha et al., 2014) and in erythroid cell lines (Pimkin et al., 2014; Stonestrom et al., 2015; Welch et al., 2004). A generalized repression of RNA production occurs during terminal erythroid differentiation, coinciding with the timing of hemoglobin synthesis. One such RNA-sequencing study of fluorescence-activated cell sorting (FACS)-enriched human erythroblast populations revealed 9,606 genes expressed in proerythroblasts, compared to 4,804 genes in orthochromatic erythroblasts (An et al., 2014). Of note, about 2,000 of these genes increased in expression over the course of erythroid differentiation. These changes in

gene expression are facilitated by alterations of the transcriptional network. GATA1 and TAL1 are "master" erythroid transcriptional factors that activate essentially all erythroid-specific genes and silence those associated with a less mature, proliferative state. Genes that are most highly induced during maturation are those related to hemoglobin synthesis, structure, and function, including the globin genes, δ-aminolevulinate synthase 2 (*ALAS2*), a hemoglobin stabilizing protein (*AHSP)*, and transferrin receptor (CD71) (An et al., 2014).

Cellular and nuclear morphological changes

Visible morphological changes enable categorization of erythroid progenitors into the various stages – basophilic erythroblasts, polychromatic erythroblasts, orthochromatic erythroblasts, and finally the anucleated reticulocyte. Erythroblasts gradually decrease in cell size and incrementally demonstrate less basophilia with histologic staining, as hemoglobin (stains pink) production increases and ribosomes (stain blue) are eliminated.

Nuclear condensation

During erythroid differentiation and maturation, large nuclei are actively reduced into condensed, transcriptionally inactive nuclei about $1/10^{th}$ of original volume. Efficient nuclear condensation is required for downstream enucleation (Hattangadi et al., 2014; Jayapal et al., 2010; Ji et al., 2010; Zhao et al., 2016). This process is mediated in part by epigenetic changes. Multiple studies highlighted the roles of histone deacetylation,

removal of acetyl groups from histone tails, which stabilizes chromatin and promotes heterochromatin formation. HDAC2 chemical inhibition or knockdown prevents chromatin condensation and larger nuclei in primary cells (Ji et al., 2010; Popova et al., 2009). Concordantly, c-Myc downregulation of the histone acetyltransferase Gcn5 was necessary for nuclear condensation. Ectopic expression of either c-Myc or Gcn5 inhibited condensation and increased nuclear size (Jayapal et al., 2010). These factors function inside the nucleus and promote chromatin compaction.

In addition to processes inside the nucleus, recent studies suggest that histones and nuclear proteins are actively transported into the cytoplasm to aid nuclear condensation. *Xpo7 (exportin-7)* is induced during late erythropoiesis, and shRNA-mediated, loss-of-function studies in fetal liver erythroblasts showed that *Xpo7*-deficient erythroblasts sustained larger nuclei, possibly due to retention of nuclear proteins within the nucleus (Hattangadi et al., 2014). Furthermore, Zhao et al. recently reported the presence of dynamic openings in the nuclear membrane, detected by discontinuous lamin B immunostaining during late erythropoiesis (Zhao et al., 2016). Inhibition of the openings by caspase-3 enzyme blockade or genetic disruption also resulted in increased nuclear size and prevented cytosolic histone release.

Enucleation

Orthochromatic erythroblasts exit the cell cycle and undergo enucleation, where the condensed nucleus separates into an anucleated reticulocyte (Sankaran et al., 2008b). The expelled nucleus is surrounded by a lipid bilayer membrane, contains a thin cytoplasm, and is called a pyrenocyte (E and D, 1967; Simpson and Kling, 1967). A growing body of evidence parallels enucleation with an asymmetric cell division, as several structures similar to those crucial in mitosis have been identified in enucleating erythroblasts. Additionally, early electron microscopy demonstrated that enucleation does not involve exocytosis, as a thin layer of cytoplasm and plasma membrane surround the extruded nuclei (E and D, 1967; Simpson and Kling, 1967). The fate of the pyrenocyte is nearly immediate engulfment by macrophages in the bone marrow. The pyrenocyte membrane is highly enriched with phosphatidyl-serine, allowing direct recognition by the macrophage (Yoshida et al., 2005).

The first morphological step of enucleation is polarization of the nucleus from a central to peripheral location in the cell. This process is largely mediated by microtubules. Treatment of late-stage erythroblasts with nocodozole or colchicine depolymerizes microtubules and leaves the nucleus in the middle of the cell (Konstantinidis et al., 2012). Exposure to taxol stabilizes microtubules and thickens the microtubule bundles surrounding the nucleus. Both destabilizing and stabilizing microtubules, however, inhibited enucleation (Konstantinidis et al., 2012). While microtubules play a structural role in the physical movements of the nucleus, they also

regulate local concentrations of secondary messengers to mediate polarization. Phosphoinositide 3-kinase (PI3K) is an important kinase which activates secondary messengers phosphatidyl-inositol(3,4)P_2 (PIP_2) and phosphatidyl-inositol(3,4,5)P_3 (PIP_3) in migrating cells, and was found to localize from the plasma membrane, directionally away from the centrosome after nuclear polarization occurred (Wang et al., 2012). Nuclear polarization was also reversed with the PI3K inhibitor, LY294002, and nocodozole treatment. Together, these studies provided evidence that microtubules directly and indirectly stabilized the nucleus by the centrosome prior to the onset of enucleation.

After the nucleus becomes polarized through microtubule-dependent processes, it is extruded away from the centrosome. At this point, the cell is no longer rounded and maintains a doublet morphology. Nuclear extrusion occurs through the formation of a contractile actin ring (CAR) and a cleavage furrow at the point between the incipient reticulocyte and pyrenocyte. Immunofluorescence experiments and time-lapse imaging of enucleating erythroblasts revealed a phalloidin-positive focus between this junction (Ji et al., 2008; Konstantinidis et al., 2012; ST and MC, 1989; Wang et al., 2012). Formation of the CAR is driven by actin polymerization and is required for enucleation, as the actin-depolymerizing agent cytochalasin D effectively inhibited nuclear extrusion (Konstantinidis et al., 2012; Ubukawa et al., 2012). Similar to the progression of mitosis in other cells, the CAR is composed of actin filaments, and formation is driven by RacGTPases (Ji et al., 2008). The final separation of the pyrenocyte and reticulocyte

occurs through a telophase-like abscission event. Lipid raft-containing vesicles deliver membrane to the abscission point, mediated in part by clathrin. Knockdown or chemical inhibition of clathrin mildly inhibited enucleation of primary human erythroblasts *in vitro* (Keerthivasan et al., 2010).

Pinpointing the functions of a protein that acts both during terminal cell divisions and enucleation can be limited because several events coordinated during enucleation utilize generalized cytokinetic machinery. These issues are potentiated by differences between *in vitro* primary cell culture systems and *in vivo* experimental mouse models. For example, mDia2 is a formin and nucleates unbranched actin filaments. Initial knockdown experiments in fetal liver erythroblast cultures suggested it specifically functions to pomote erythroid nuclear extrusion (Ji et al., 2008). However, Watanabe et al. showed that mDia2 function is highly intertwined with the prior cell divisions, as mDia2 knockout mice have increased multinucleated erythroid progenitors (Watanabe et al., 2013). Importantly, enucleation is the final step of erythroid development, and defects in numerous critical processes upstream can result in impaired enucleation.

Cell cycle: regulation of RBC size and number

Regulation of the cell cycle is intimately correlated with the cell number and size produced at the end of terminal differentiation. The earliest progenitors undergo specialized cell divisions producing two daughter cells, each smaller in size than the

parental cell. The G1 phase is shortened and S phase becomes progressively shorter, both allowing for decreased size after mitosis (Dolznig et al., 1995; Grebien et al., 2005). Cyclin D3 and cyclin A2 promote the G1/S and G2/M transitions of the cell cycle, respectively, and single nucleotide polymorphisms (SNPs) were identified by genome-wide association studies (GWAS) to be associated with deviations in mean corpuscular volume and RBC count (cyclin D3) and cell size (cyclin A2) (Ganesh et al., 2009; Kamatani et al., 2010; van der Harst et al., 2012). Indeed, follow-up loss-of-function studies validated roles for these cyclins in erythroid maturation in primary murine erythroid culture systems (Ludwig et al., 2015; Sankaran et al., 2012). Erythroblasts expressing cyclin D3-targeting shRNAs displayed fewer cell divisions during the maturation process, resulting in larger and fewer resultant RBCs (Sankaran et al., 2012). Meanwhile, cyclin A2 knockdown in fetal liver erythroblasts inhibited cell size by regulating cytokinesis during the last cell division (Ludwig et al., 2015). In addition, these studies support the idea that successful production of functional erythrocytes is independent of abnormal cell size.

Perturbation of the specialized divisions during erythroid maturation is an underlying cause of a rare condition called congenital dyserythropoietic anemia, wherein the pathological hallmark of the disease is multinucleated erythroblasts in the bone marrow. Causal mutations have been identified in CODANIN I (*CDAN1*) (Noy-Lotan et al., 2009; Renella et al., 2011), *SEC23B* (Schwarz et al., 2009), and most recently *MKLP1* (*KIF23*) (Liljeholm et al., 2013). While all these genes are ubiquitously

expressed, the disease is largely erythroid-restricted, and studies of these mutations have expanded our understanding of the cell divisions during terminal erythropoiesis. CDANI remains a poorly understood protein that may facilitate histone assembly into chromatin, and *SEC23B* encodes a protein involved in endoplasmic reticulum vesicle trafficking. Meanwhile, *MKLP1* is a well-studied kinesin-like motor protein part of the centralspindlin complex, which organizes microtubules (Mishima et al., 2002), regulates formation of the CAR (Pavicic-Kaltenbrunner et al., 2007), connects the central spindle to the plasma membrane in preparation for abscission (Lekomtsev et al., 2012). How mutations in these three ubiquitously expressed proteins are responsible for an erythroid-restricted disease is still unclear.

1.4 The Ubiquitin-proteasome system (UPS)

The UPS facilitates selective elimination of proteins, including those that are irreversibly damaged, cytotoxic, or unnecessary. Three cooperating proteins coordinate the ATP-dependent ubiquitination of the target protein. An E1 enzyme activates the 76-amino acid polypeptide ubiquitin and transfers it to an E2 ubiquitin-conjugating enzyme. E3 ubiquitin ligases mediate direct or indirect covalent linkage of ubiquitin monomers or chains to substrates (Hershko et al., 1983). After the first ubiquitin is attached, ubiquitin moieties are added to a lysine (K) reside of the previously attached ubiquitin: K6, K11, K27, K29, K33, K48, or K63). K48-linked polyubiquitination classically, but not exclusively, mediates recognition by the proteasome, a multi-subunit organelle that

includes numerous proteases (Thrower et al., 2000). However, it is now known that ubiquitin attachments are post-translational modifications serving innumerable functions, including cell cycle regulation, intracellular transport, cell signaling, and DNA repair (Erpapazoglou et al., 2012; Spence et al., 1995; Xu et al., 2009).

While there are <10 E1 and ~40 E2 enzymes in cells, the diversity of E3 ubiquitin ligases is much greater; over 600 are expressed (Li et al., 2008). Each recognizes a limited repertoire of substrates and bestows specificity to the UPS. E3 enzymes are categorized into two main families, depending on the presence of a HECT (Homologous to E6-AP Carboxyl Terminus) or a RING (Really Interesting New Gene) domain. While HECT-containing enzymes carry activated ubiquitin and directly ligate the group to substrate, RING-containing enzymes act as scaffolds to bind the E2-ubiquitin enzyme and substrate to indirectly mediate ubiquitination (Deshaies and Joazeiro, 2009).

The UPS in erythropoiesis

The system was initially described in reticulocytes (Ciechanover et al., 2012; Etlinger and Goldberg, 1977; Wilkinson et al., 1980). Gene set enrichment analysis of the global transcriptome of differentiating human erythroblasts revealed that UPS components are highly enriched during late erythropoiesis (An et al., 2014; Egan et al., 2015). In parallel, proteomics revealed that about 1/3 of known molecules of ubiquitination pathways exist in the human RBC (Roux-Dalvai et al., 2008). Evidence of

the broad role of the UPS in erythropoiesis also stems from the use of small molecule proteasome inhibitors. Studies using cultured erythroid cells demonstrated that exposure to various proteasome inhibitors can dramatically inhibit erythroid differentiation and maturation (Chen et al., 2002; Wölwer et al., 2015). A chemical screen on cultured, maturing erythroblasts identified three proteasome-inhibiting compounds that inhibited enucleation (Wölwer et al., 2015). One of these, bortezimib, is currently prescribed to treat multiple myeloma, and an observed side effect of its use is anemia (San Miguel et al., 2006). Despite having these clues, relatively few studies have elucidated substrates recognized and assigned functions to UPS components during red blood cell formation.

Trim proteins

The Trim protein family consists of more than 70 members, defined by a tripartite motif containing RING, B box, and coiled-coil (CC) domains (Marín, 2012). The canonical amino acid motif in RING domains is two repeats of Cys-X-His-X_2-Cys-X_2-Cys, where X is any amino acid (Freemont et al., 1991). The conserved cysteine and histidine residues bind two atoms of zinc, maintain the overall protein structure, and form a cleft on the surface of the RING domain, which supports binding of the E2 protein (Deshaies and Joazeiro, 2009; Zheng et al., 2000).

The CC domain enables Trim protein oligomerization. The crystal structures of the CC domains of TRIM5a, TRIM20, TRIM25, and TRIM69 all demonstrate that CC

domains form interdigitating antiparallel hairpins, thereby positioning the RING domain of one oligomer directly apposed with the C-terminal substrate-binding domain (Goldstone et al., 2014; Li et al., 2014; Sanchez et al., 2014; Weinert et al., 2015). Furthermore, several recent studies indicate that Trim protein multimerization is critical for efficient enzymatic activity. Full-length TRIM25 and TRIM32 proteins successfully formed dimers in solution and showed autoubiquitination activity, whereas truncated mutant proteins lacking the CC domains were attenuated in both *in vitro* assays (Koliopoulos et al., 2016; Streich et al., 2013). Catalytic activity of TRIM32 was partially rescued by fusion of the RING domain to glutathione-S-transferase domain, which is known to mediate dimerization (Streich et al., 2013).

Trim proteins have variable C-terminal domains that commonly mediate protein interactions. More than half contain a C-terminal PRY-SPRY (Sp1a and Ryanodine Receptor) domain (Reymond et al., 2001; Versteeg et al., 2013)2. Crystal structures of other PRY-SPRY domains reveal six hypervariable immunoglobulin-like loops that interact with substrate peptide (Biris et al., 2012; Woo et al., 2006) (D'Cruz et al., 2013) (James et al., 2007). Of note, mutations in the PRY-SPRY domain of TRIM20, or PYRIN, have been linked to an autoimmune disorder named Familial Mediterranean Fever (Grütter et al., 2006; Weinert et al., 2015; Woo et al., 2006).

1.5 β-globin gene regulation

As discussed in Section 1.3.1, terminally differentiating and maturing erythroblasts utilize several mechanisms to dedicate a vast majority of cellular and transcriptional machinery to produce hemoglobin.

Historically, scientists debated whether replacement of HbF by HbA was caused by a switch in stem cell lineage between fetal development and adulthood. Does an RBC expressing HbF have a different origin than an RBC expressing HbA? A series of clonal analyses of primary erythroid progenitors revealed that erythroid colonies were comprised of distinct populations of cells expressing HbA and cells expressing HbF (Papayannopoulou et al., 1976; 1977). These studies not only indicated that adult and fetal erythroid cells indeed share a common stem cell lineage, but also supported a model where transcription network changes were responsible for hemoglobin switching. The molecular mysteries of this transcriptional switch have spurred intense basic research with the ultimate goal to discover new, targeted therapies for hemoglobinopathies.

The β-like globin genes

The five β-like globin genes lie on human chromosome 11 in the following order: 5'-ε-Gγ-Aγ-δ-β-3' (Figure 1.2). ε is expressed during embryogenesis, Gγ and Aγ are expressed during fetal development, and δ-globin and β-globin replace the fetal genes within 6 months after birth. Upstream of the β-globin genes lies a cluster of five

conserved DNase I hypersensitive sites (HSs), comprising the locus control region (LCR). The HSs function as erythroid-specific enhancers and mediate high-level induction of the β-like globin genes in maturing erythroid cells ((Forrester et al., 1986; Kollias et al., 1986; Tuan et al., 1985), reviewed in (Li et al., 2002)). The absence of the LCR dramatically reduces β-globin expression (Alami et al., 2000; Bender et al., 2000). Lying 20-60 kb upstream of the β-globin genes, the LCR remotely activates gene expression by chromosomal looping. When the embryonic genes are expressed during primitive erythropoiesis, the LCR interacts with the embryonic globin genes and not the adult globin genes. The LCR interactions with the gene promoters switch to temporally coordinate the expression of the different globins (Carter et al., 2002; Tolhuis et al., 2002).

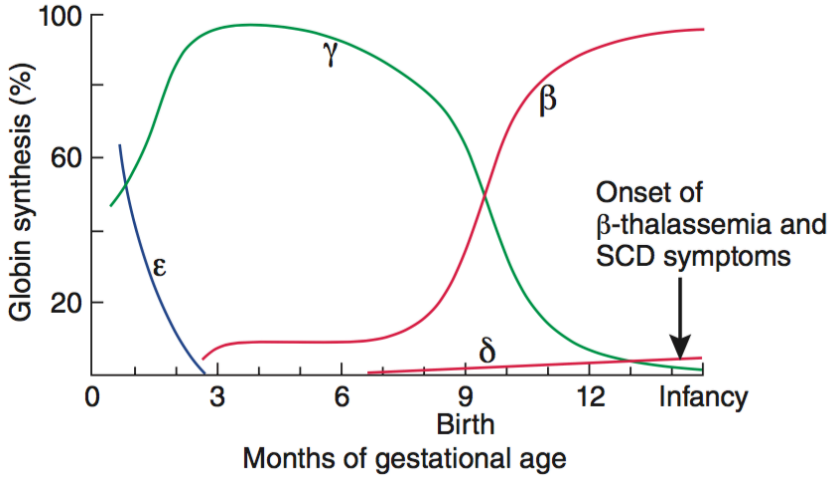


Figure 1.2. Timing of the β-like globin gene switching during human development. The human embryonic, fetal, and adult globin chain relative expression levels are shown in blue, green, and red, respectively. From (Sankaran and Weiss, 2015).

Sickle cell disease

Sickle cell disease (SCD) was the first "molecular disease" and is caused by a substitution of valine for glutamic acid, the 6th amino acid (D6V) (Pauling et al., 1949). The mutant protein is incorporated into hemoglobin tetramers, which polymerize under deoxygenating conditions and causes the RBCs to transform from a biconcave disk to a sickle shape (HerrickIrons, 1910). Vaso-occlusive events are precipitated when the cells enter smaller capillaries or the microvasculature and become deoxygenated, producing elongated cells that obscure blood flow to the tissue. These complications include stroke, pain, acute vascular necrosis, nephropathy, and ultimately premature mortality.

Rationale for HbF-inducing therapies

Studies of the interactions between HbS and other forms of hemoglobin has provided the rationale for new therapies (Behe and Englander, 1979; Sunshine et al., 1979). When mixed with HbS, HbA and HbF both significantly delayed and decreased polymerization of HbS. However, HbF acts more potently. Compared to the polymerization time of pure HbS, mixtures with 30% HbA have delayed polymerization by 100-fold. However, when HbS is mixed with 30% HbF, polymerization is 10^4-fold slower. Therefore, the biochemical effects of having 30% HbF tetramers are more beneficial than observed the same ratio of HbA. In parallel, studying patients with sickle

cell disease and varied diseased progressions have supported similar conclusions that HbF expression alongside HbS is therapeutically beneficial. Baseline values of HbF amongst sickle patients is variable, ranging from <5% to >30%. In the Cooperative Study Sickle Cell Disease, HbF above 4% was associated with reduced painful crises (Platt et al., 1991). A slightly greater increase in HbF level above 8.6% was associated with decreased mortality (Platt et al., 1994). To date, this is the only known genetic modifier of sickle cell disease death.

BCL11A

Bcl11a is a zinc-finger transcription factor historically known for its role in lymphocyte and neural development. Because HbF expression and fraction of HbF^{+} cells (F-cells) are quantitative traits, the dawn of modern genomics has opened the doors of discovery. Genome-wide association studies elucidated the association between common variants in the *BCL11A* locus on chromosome 2 with fetal hemoglobin levels (Menzel et al., 2007; Uda et al., 2008). Knockdown of *BCL11a* in primary human erythroblasts cells showed that it is a negative regulator of γ-globin expression (Sankaran et al., 2008a), and erythroid-specific knockout mice inappropriately expressed γ-globin into adulthood (Xu et al., 2011). Furthermore, co-inheritance of Bcl11a null alleles rescued the disease manifestations of SCD in a humanized mouse model (Xu et al., 2011). Genetic studies have identified rare occurrences of large genomic deletions including the *BCL11A* gene. While these individuals have increased HbF expression, these deletions are also associated with neurocognitive and developmental disorders (Basak et al., 2015; Funnell

et al., 2015). Therefore, global or total-body inhibition of BCL11A by an ingested drug, for instance, may have adverse effects on the brain.

More recently, Bauer et al. described an erythroid-specific enhancer of *BCL11A* (Bauer et al., 2013). They used previous GWAS data to fine-map a region of intron 2 in *BCL11A*. By matching the GWAS variants with DNase I hypersensitive sites found in erythroid cells (but not in other cell types), three HS regions (+55, +58, and +62 kb from the transcription start site) corresponded to SNPs linked to HbF variation. Deletion of the enhancer in cell lines left the exonic sequence intact, but loss of the enhancer reduced expression of *BCL11A*. The group further probed the 12-kb enhancer sequence to identify the minimal region necessary for enhancer function by performing a saturating CRISPR/Cas9 mutagenesis screen (Canver et al., 2015). Mutations induced by a single gRNA targeting the +58 HS site upregulated γ-globin to the same level as deletion of the entire enhancer. In parallel, Vierstra et al. screened eight zinc finger nuclease pairs targeting the enhancer and found that the pair targeting within 20 nt of +58 was the most effective at raising HbF expression (Vierstra et al., 2015). Both studies suggest that enhancer activity, *BCL11A* expression, and subsequent HbF repression are dependent on a single GATA1 motif at +58. The strong, safe effects of variation at this locus and target resolution have made this locus the prime candidate for therapeutic genome editing.

1.6 HPFH

Individuals heterozygous or homozygous for mutations that cause HbF expression are asymptomatic. Usually identified incidentally by routine screening or because other family members have a hematologic disorder, rare individuals have mutations in the β-globin gene cluster that cause hereditary persistence of fetal hemoglobin (HPFH). These individuals have normal hemoglobin concentrations.

Several different models are currently thought to explain mechanisms of HbF induction seen with HPFH. 1) Enhancer sequences are brought closer to *HBG1* and *HBG2* (Forget, 1998)*;* 2) removal of *HBB* and *HBD* eliminate competition with the LCR (Akinsheye et al., 2011); and 3) regulatory sequences that play a role in γ-globin gene repression are altered or created *de novo* (Wienert et al., 2015).

Deletional HPFH

Several deletions in the β-globin gene cluster result in HPFH. Heterozygotes have reported greater than 30% HbF in a pancellular distribution, where all RBCs express HbF. Homozygotes are clinically healthy but may have mild microcytosis (smaller RBCs) and/or hypochromia (less hemoglobinization). The α-globin to γ-globin chain synthesis can be mildly imbalanced, suggesting that γ-globin synthesis occurs at a lower rate than the normal β-globin gene does. A schematic of the known deletions is shown in Figure 1.2 (drawn based on references (Feingold and Forget, 1989; Henthorn et al., 1990;

Huisman et al., 1971; Jagadeeswaran et al., 1982; Joly et al., 2009; Sankaran et al., 2011)). HPFH-1 and HPFH-2 mutations are the most common and are large deletions (>84 kb in length) and often encompass the *HBB* gene itself (Feingold and Forget, 1989). These mutations genetically resemble and even overlap with those responsible for δβ-thalassemia. However, the mapping of these breakpoints highlighted a 3.5-kb region within the β-globin gene cluster that is critical for γ-globin silencing (Sankaran et al., 2011).

Figure 1.3. Map of known HPFH deletions in the β-globin gene cluster on chromosome 11.

Vertical arrows indicate the HS sites of the LCR, boxes show the β-like genes, and horizontal arrows denote the span of the deletions. Elements are drawn to scale.

13-nucleotide HPFH deletion

Huisman et al. reported the smallest known HPFH deletion: a 13-nucleotide (nt) mutation in the *HBG1* promoter, 102 through 114 nt upstream of the transcriptional start site (-102 to -114) (Gilman et al., 1988). The mutation was identified in two related individuals with sickle cell trait and HbF levels of 31.8% and 30.1%. Aγ, encoded by *HBG1*, represented over 80% of the γ-globin chains, suggesting that the mutation in the promoter was responsible for the specific upregulation of Aγ−globin. Of note, three additional point mutations lie in the same region, including a G-to-A change at -117 (-117G>A, 10-20% HbF) (Berry et al., 1992; Collins et al., 1985), a C-to-T change at -114 (-114C>T, 11-14% HbF) (Fucharoen et al., 1990), and a G-to-T change at -109 (-109G>T, 4.1% HbF) (Chassanidis et al., 2009). Transgenic mouse experiments served to validate that these mutations were responsible for HPFH phenotypes. While a transgenic mouse model is lacking for the 13-nt deletion, Ronchi et al. demonstrated that the -114 and -117 point mutations indeed induced γ-globin expression both in fetal liver and peripheral blood of adult animals (Ronchi et al., 1996). This study also highlighted the necessity for intact proximal elements upstream of -114. The γ-globin promoter mutations provided the backbone for *in vivo* structure-function mapping and suggested that these nucleotides are specifically important for hemoglobin switching.

Chapter 2 Experimental Methods

Trim58 cloning

The full-length Trim58 gene, including 5' and 3' untranslated regions (UTR), was initially TOPO cloned into pCRII-Dual vector (Invitrogen) from murine fetal liver erythroblast cDNA using the following primers (Forward: GCCATGGCCACGGCACCCGGGG, Reverse: CAGCACTTCTGGATGGGTTT). Full length (wild type) Trim58 was subcloned from pCRII-Dual-Trim58 into retroviral vectors using the following primers to create 5' FLAG/HA epitope-tagged constructs (Forward: GAAGATCTAGATCTGCCATGGACTACAAGGACGACGATGACAAATACCCATA CGACGTCCCAGACTACGCTGCCTCAGCTCCTTCTGTG, Reverse: CGGAATTCCAGCACTTCTGGATGGGT). The RING-dead Trim58 construct, containing Cys>Ala missense mutations at residues 55 and 58 (within the RING domain), was created from a full length Trim58 construct using a two-step PCR method using the above full length Trim58 Forward and Reverse primers and the following mutation-generating primers (Forward: GGACCCCTGGCCTGGGGTGCGGCGTAGAC, Reverse: GTCTACGCCGCACCCCAGGCCAGGGGTCC). The Trim58 PRY-SPRY domain, including residues 281-485, was subcloned from a full length Trim58 construct using the following primers (Forward: AGAAGATCTAGATCTGCCATGAGGGAGATG, Reverse: CGGAATTCCAGCACTTCTGGATGGGT). This construct begins 8 amino acids upstream and extends 25 amino acids downstream of the predicted PRY-SPRY domain.

Live cell imaging plasmids were cloned into the BglII/EcoRI site of a pK1 vector containing amino-terminal FLAG and carboxy-terminal mCherry tags (S. Kadauke & G. Blobel, Children's Hospital of Philadelphia) using the following primers (Forward: GAAGATCTAGATCTATGGCCACGGCACCCG, Reverse: CCGGAATTCGAATTCGTATTCCTCACTTCTGGCAG). *Trim58* ΔCC expressed in E14.5 fetal liver cells as isolated by PCR from a murine fetal liver erythroblast cDNA library.

Trim58 constructs were cloned into the BamHI/EcoRI or BglII/EcoRI sites in several expression vectors for this study, including pCRII-Dual (Invitrogen), retroviral MIGR1 (Murine Stem Cell Virus (MSCV) promoter-driven construct with downstream IRES-GFP), retroviral pK1 (MSCV-driven construct with downstream IRES-puromycin), and pGEX6P1 (GE Healthcare) for GST-tagged constructs.

Radioactive in situ hybridization

Embryonic day 14.5 (E14.5) murine embryos were fixed, dehydrated, paraffin-embedded, sectioned, and stained with an antisense probe targeting Trim58 nucleotides 841-1455 per protocols from the University of Pennsylvania Molecular Cardiology Research Center Histology and Gene Expression Core Facility.

Flow cytometry

For fluorescence activated cell sorting of primary murine fetal liver erythroblasts, single cell suspensions from whole fetal livers were stained with Ter119-APC

(BioLegend) and CD71-PE (BD Pharmingen) for 45 min at 4 °C in PBS and sorted on a FACSAria instrument (BD Biosciences) per a previously described gating strategy (Pop et al., 2010).

For cultured murine fetal liver erythroblast analysis, $5x10^5$ cells were stained sequentially with 5 μM Hoescht33342 (Sigma) for 1 hr at 37 °C in fetal liver maturation medium, Live/Dead near-IR fixable dead cell stain (Invitrogen) for 30 min at 4 °C in PBS, and Ter119-PerCP-Cy5.5 and CD44-AF647 (BioLegend) for 45 min at 4 °C in PBS with 2% fetal bovine serum (FBS). Cells were analyzed on an LSR Fortessa instrument (BD Biosciences) maintained by the Flow Cytometry Core Laboratory of The Children's Hospital of Philadelphia Research Institute. Data were analyzed using FlowJo software (TreeStar).

Quantitative real-time PCR

RNA was extracted using the RNeasy kit (Qiagen) with on-column DNAse treatment. cDNA was prepared using the iScript cDNA kit (BioRad). Semiquantitative real-time PCR was done using the standard curve method and SYBR green dye on a ViiA 7 real-time PCR system (Life Technologies). Real time primers used in this study probed for murine Trim58 (*Trim58*; Forward: GAGCGTCTTTGGAACTTGTG, Reverse: ACCCTCTGTGTTTCTCAAACTC), dynein heavy chain (*Dync1h1*; Forward: TTGTACCGCATCCAAGAGAAG, Reverse: GTTGTAGTCATTCACCGTTTCC), dynein intermediate chain (*Dync1i2*, Forward: ACAGTCAAAGGCAGTAGCTG,

Reverse: CTGGTGTCCCTCAAACATCTC), actin (*Actb*; Forward: CCTTCCTTCTTGGGTATGGAATC, Reverse: AGCACTGTGTTGGCATAGAGGT), hypoxanthine guanine phosphoribosyl transferase (*Hprt*; Forward: TCAGTCAACGGGGGACATAAA, Reverse: GGGGCTGTACTGCTTAACCAG), and glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*, Forward: AGGTTGTCTCCTGCGACTTCA, Reverse: CCAGGAAATGAGCTTGACAAA). Target gene expression was normalized to the average of *Actb, Gapdh,* and *Hprt* values. *HBG1/2* and *HBB* mRNAs were quantified by SYBRGreen qPCR, as described (Deng et al., 2014).

In silico **analyses of Trim58 gene regulation and expression**

ChIP-Seq data at the murine *Trim58* locus were obtained from the Penn State University Bioinformatics Genome Browser (https://mery.genome-browser.bx.psu.edu/). The data presented in Figures 3.2C and 3.23 were generated from primary fetal liver erythroblasts by ChIP-Sequencing and RNA-Sequencing, respectively (Pimkin *et al,* in revision). Human *TRIM58* mRNA tissue expression patterns were obtained from publicly available microarray data (Figures 3.1A (Wu et al., 2009) and 3.1B (Novershtern et al., 2011)). Erythroid *TRIM58* expression data were obtained from microarray analysis of FACS-purified human erythroblasts cultured from peripheral blood buffy coat mononucleocytes (Figure 3.1C (Merryweather-Clarke et al., 2011)).

Murine fetal liver erythroblast isolation

E14.5 murine fetal livers were harvested from pregnant CD1 mice (Charles River Laboratories), triturated, and passed through a cell strainer (BD Biosciences) to yield stroma-free single cell suspensions. Erythroid progenitors were purified by negative selection on magnetic beads using the EasySep Mouse Hematopoietic Progenitor Cell Enrichment Kit (STEMCELL Technologies) according to the manufacturer's instructions, with the addition of biotin-conjugated anti-mCD71 (eBioscience).

Short hairpin RNA cloning

Short hairpin RNAs from pGIPZ vectors (Open Biosystems) were subcloned into the "PIG" (MSCV-Puromycin-IRES-GFP) retroviral vector (Hemann et al., 2003). These shRNA constructs are within a miR30 stem-loop structure to aid processing. Primers (Forward: AGATCTAGATCTTGCTGTTGACAGTGAGCG, Reverse: CTCGAGCTCGAGTCCGAGGCAGTAGGC) were used to clone the following shRNAs into the MSCV-PIG vector BglII and XhoI sites. The sense strand is indicated in red, antisense in blue, loop in green, and the common miR-30 context in black.

shLuciferase (RHS1705)
TGCTGTTGACAGTGAGCGCCCGCCTGAAGTCTCTGATTAATAGTGAAGCCAC
AGATGTATTAATCAGAGACTTCAGGCGGTTGCCTACTGCCTCGGA

shScrambled (RHS4346)

TGCTGTTGACAGTGAGCGATCTCGCTTGGGCGAGAGTAAGTAGTGAAGCCAC
AGATGTACTTACTCTCGCCCAAGCGAGAGTGCCTACTGCCTCGGA

shTrim58 #3 (V3LMM_487333)

TGCTGTTGACAGTGAGCGACCGAGACTCAGTGCTGAGAAATAGTGAAGCCAC
AGATGTATTTCTCAGCACTGAGTCTCGGCTGCCTACTGCCTCGGA

shTrim58 #4 (V3LMM_487334)

TGCTGTTGACAGTGAGCGCCTCCTTCTACAATGTCACAAATAGTGAAGCCAC
AGATGTATTTGTGACATTGTAGAAGGAGATGCCTACTGCCTCGGA

shTrim58 #5 (V3LMM_487335)

TGCTGTTGACAGTGAGCGCTTGGACTACGAAGCTGGTGAATAGTGAAGCCAC
AGATGTATTCACCAGCTTCGTAGTCCAAATGCCTACTGCCTCGGA

shTrim58 #7 (V3LMM_487337)

TGCTGTTGACAGTGAGCGCCGGGATCTTTTTGGACTACGATAGTGAAGCCAC
AGATGTATCGTAGTCCAAAAAGATCCCGATGCCTACTGCCTCGGA

Retroviral infection

Retroviral PIG constructs were packaged by co-transfection of pCL-Eco into 293T cells via calcium phosphate transfection. To create retrovirus for human (HeLa) cell infection, constructs were packaged by co-transfection of pVSVG and pCPG into 293T cells. Retroviral supernatants were collected 24 and 48 hr post-transfection and kept at -80 ^{o}C until use. Cells were infected with retroviral supernatant containing 8 μg/mL polybrene (Sigma) and 10 mM Hepes (Gibco) by centrifugation at 2800 rpm for 90 min at 30 °C.

Fetal livery erythroblast culture conditions

Our methods were based on those initially devised by Zhang et al (Zhang et al., 2003) with some modifications. Erythroid progenitors were “expanded” for 24-72 hrs in StemPro-34 serum free media (Gibco) containing 10% SP34 supplement, 2 mM L-

glutamine (Gibco), 1% Penicillin/Streptomycin (Gibco), 0.1 mM 1-thioglycerol (Sigma), 1 μM dexamethasone (Sigma), 0.5 Units/mL Erythropoietin (Amgen), and 1% mSCF-conditioned medium. At 18-24 hrs post-infection, puromycin (1 μg/mL) was added to select for infected cells. Following expansion, cells were pelleted at 800 rpm for 3 min, washed thrice in PBS, and cultured in fetal liver maturation medium (IMDM (Mediatech/Cellgro), 10% BenchMark fetal bovine serum (Gemini Bioproducts), 10% plasma derived serum (Animal Technologies), 5% Protein Free Hybridoma Media II (Gibco), 2 mM L-glutamine (Gibco), 1% Pen/Strep (Gibco), 0.1 mM 1-thioglycerol (Sigma), and 2 Units/mL Erythropoietin (Amgen)) for up to 48 hrs. After 48 hrs, maturation was complete, cultures began to die, and analyses became unreliable. Expansion and maturation cultures were kept below 1 million cells per mL at all times.

Erythroblast morphology analysis

Cells were centrifuged onto a glass slide and stained with May Grünwald-Giemsa (Sigma). Light microscopy images were obtained with a Zeiss Axioskope 2 microscope, Zeiss Axiocam camera, and Zeiss AxioVision 3.1 software (Carl Zeiss Microimaging) at room temperature. Multinuclearity was assessed by visual inspection of May Grünwald-Giemsa-stained slides. The “Analyze Particles” feature within FIJI was used to quantify nuclear sizes from slide images.

Western blot analysis

Sample loading was normalized to protein content. Proteins were resolved on Tris-Glycine gels (BioRad), transferred to 0.45 μm PVDF (Whatman) or Nitrocellulose membrane (BioRad), and blocked in 5% milk. Primary antibodies included Dynein Heavy Chain (R-325, Santa Cruz), Dynein Intermediate Chain (MAB #1618, Millipore), Nup153 (QE5, Abcam), Gapdh (FL-335, Santa Cruz), α-globin (custom rabbit polyclonal, Covance), Hemagglutinin (Y-11, Santa Cruz), FLAG (M2, Sigma), Codanin1 (G-1, Santa Cruz), and β actin-HRP (Sigma). Anti-Trim58 antibodies were raised against a peptide from the N-terminus (ERLQEEARCSVCLDFLQEPISVD) of murine Trim58 (Thermo Scientific). HRP-conjugated anti-mouse and anti-rabbit secondary antibodies, protein size markers, and other reagents were from Thermo Scientific.

Hemoglobin content quantification

Hemoglobin content in mature erythroblast cultures was quantified by Drabkin's assay as described (Campbell et al., 2013). Culture pellets were imaged in 1.1 mL FACS tubes (VWR) with an iPhone 4.

Immunoprecipitation assays

For all immunoprecipitation (IP) experiments, cells were lysed in buffer containing 10 mM Tris (pH 7.4), 150 mM sodium chloride, 0.5% NP-40 (Sigma), 1 mM EDTA (pH 8.0), 10 μM proteasome inhibitor (MG132, Enzo), and 1:500 protease inhibitor cocktail (Sigma). For FLAG- PS immunoprecipitation, 400 million FACS-purified GFP+ G1E cells, cultured as described (Weiss et al., 1997), were lysed,

precleared with a 1:1 mix of Protein A and rProtein G agarose (Invitrogen), and incubated overnight with EZView Red Anti-Flag M2 Affinity Gel (Sigma) at 4 ^{o}C. After washing thrice with IP buffer, immunoprecipitated material was eluted off of beads with 100 μg/mL 3X FLAG peptide (Sigma) at 4 °C for 1 hr with intermittent agitation and used for further analyses.

For DIC immunoprecipitation, precleared lysate from 15 freshly isolated whole E14.5 murine fetal livers (~300 million cells) were incubated with anti-DIC antibody (Millipore MAB #1618) overnight. rProtein G agarose beads (Invitrogen) were then added for 4 hrs at 4 °C. Beads containing immunoprecipitated material were washed thrice in IP buffer, boiled in 2X Laemmli Sample Buffer (Sigma), and analyzed immediately.

Mass spectrometry

Immunoprecipitates were size-fractionated by SDS-PAGE (4-15% gradient gel, BioRad) and stained with Coomassie Blue Silver overnight. Following washing in deionized water, the indicated bands were manually extracted, digested with trypsin, and subjected to liquid chromatography and nanospray/linear trap quadrupole mass spectrometry using a ThermoFinnigan LTQ linear ion trap mass spectrometer at the University of Pennsylvania Proteomics Core Facility. Data were analyzed via Sequest and Scaffold3 software packages. Presented data represent unweighted spectral counts, unique peptide counts, and percent protein coverage for proteins identified by ≥2 peptides with >99% protein/>95% peptide confidence.

In vitro GST pull down assays

For the *in vitro* binding experiments shown in Figures 3.10A and 3.10B, GST and GST-PS proteins were prepared in previously described conditions (Kihm et al., 2002). *E. coli* BL21 cells were grown at 37 °C for 3 hrs and induced to express protein by addition of 0.1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 3 hrs at 30 °C. Cell pellets were resuspended in BC500 buffer (10 mM Tris-HCl (pH 8.0), 500 mM KCl, 20% glycerol, 1% NP-40, 0.5 mM EDTA, 1 mM DTT, 0.5 mg/mL lysozyme, 0.05 mg/mL DNAse I, and protease inhibitor cocktail (Sigma)) and sonicated. Lysates were clarified by centrifugation at 10,000 rpm for 10 min at 4 °C. Cleared lysates were incubated with washed glutathione-Sepharose beads (GE Health Sciences) overnight at 4 °C, washed thrice in BC500, and eluted off of beads with 20 mM reduced glutathione solution containing 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 1 mM DTT, and protease inhibitor cocktail (Sigma).

Bovine brain holodynein was purified as described (Bingham et al., 1998). pCMV-based HA-tagged portions of human dynein intermediate chain (generously donated by K. Kevin Pfister, University of Virginia) were introduced into 293T cells by calcium phosphate transfection. Transfected 293T cells expressing HA-tagged DIC constructs were lysed in pull down (PD) Buffer (20 mM Tris pH 7.4, 150 mM potassium chloride, 1 mg/mL BSA, 1:500 protease inhibitor cocktail (Sigma)) 24 hrs after transfection.

For pull down assays, recombinant GST or GST-PS (1 μmol at 4 mM final concentration) was incubated with substrate (5 μg holodynein (~4 fmol at 16 nM final concentration) or 400 μg 293T lysate containing DIC truncation) and 20 μL glutathione-Sepharose beads (GE Healthcare) in PD Buffer for 2 hrs at 4 °C with rotation. All pipette tips and Eppendorf tubes were passivized with 1 mg/mL bovine serum albumin (Sigma) prior to use. Beads were subsequently pelleted, washed thrice with PD Buffer, boiled in 2X Laemmli Sample Buffer (Sigma), and analyzed.

SEC-MALLS experiments

Our collaborators, Drs. Joel Mackay and Ana Silva, of the School of Molecular Bioscience at the University of Sydney, Australia, conducted these experiments. Trim58 PS (residues 281-485) and dynein intermediate chain (residues 1-120) were expressed separately in *E. coli* Rosetta 2 (DE3) cells at 25 °C for 16 hrs. Cell pellets were resuspended in 50 mM sodium phosphate (pH 7.5) and 0.5 M sodium chloride for GST-PS, or 50 mM CHES pH 10 and 1 M sodium chloride for GST-DIC(1-120). Cells were lysed by sonication following the addition of 1 mg/mL lysozyme, 1 mM DTT, 1 mM PMSF, 0.05 mg/mL DNAse I, 0.05 mg/mL RNAse A and Complete EDTA-free protease inhibitors (Roche). Cell lysates were incubated with glutathione-Sepharose resin (Novagen) for 1 hr at 4 °C and washed. GST-DIC(1-120) was eluted using 30 mM glutathione, whereas PS was cleaved from the GST tag with HRV-3C protease overnight at 4 °C. Both proteins were then purified by size exclusion chromatography, using a Superdex 75 HiLoad 16/600 column. PS was eluted in 20 mM sodium phosphate (pH

7.5) and 50 mM sodium chloride, and GST-DIC(1-120) in 20 mM CHES pH 10 and 100 mM sodium chloride.

Purified PS (200 μL at 3.3 μM) and GST-DIC(1-120) (200 μL at 16.4 μM) were run separately on a Superdex 200 10/300 GL column connected to a MiniDawn Treos (Wyatt Technology). A mixture of PS and GST-DIC (200 μL of a mixture of PS at 13.2 μM and GST-DIC at 16.4 μM) was incubated for 1 hr at room temperature and then eluted on the same SEC-MALLS setup. All samples were eluted in 20 mM Tris pH 8 and 100 mM sodium chloride at 0.5 mL/min. Molecular masses were calculated using a dn/dc value of 0.185 mL/g.

Immunofluorescence assay sample preparation

HeLa cells were transiently transfected on 12 mm coverslips (VWR) with pK1-based Trim58 constructs or pEGFP-based mCherry-CC1 (human), using Lipofectamine 2000 reagent (Invitrogen) per the manufacturer's instructions. After 36 hrs, cells were fixed in 4% paraformaldehyde for 5 min, permeabilized in 0.1% Triton-X 100 for 5 min, and blocked in PBS containing 3% FBS for 1 hr at room temperature. Transfected HeLa cells were stained with GM130 (BD Biosciences) overnight at 4 °C, and then with goat anti-mouse Alexa Fluor 488 (Invitrogen) for 1 hr at room temperature. Cells were mounted on glass slides with Prolong Gold Antifade Reagent with DAPI (Invitrogen).

For erythroblast immunofluorescence assays, cells were layered onto glass slides previously treated with Vectabond reagent (Vector Laboratories) and incubated for 10

min at 37 °C. Erythroblasts were fixed and permeabilized as described above, stained with Alexa Fluor 488-conjugated anti-aTubulin (Millipore) overnight at 4 °C, and mounted.

Time lapse microscopy

HeLa cell lines were plated on a FluoroDish (World Precision Instruments, 0.17 mm-thick glass bottom) in DMEM (Gibco) containing 10% FBS and 1% sodium pyruvate at a concentration of 250,000-500,000 cells per mL. Beginning 24-36 hrs later, cells were imaged at 5 min intervals for 12 hrs in an environmental chamber at 37 °C and 5% carbon dioxide. Time lapse and static erythroblast imaging were performed on an Olympus IX70 inverted microscope with a Photometrics CoolSnap HQ high-resolution CCD camera. Images were acquired and processed using Deltavision Softworx software. Static HeLa cell imaging was performed on a Zeiss LSM 710 Confocal microscope with ZEN 2011 acquisition software. The University of Pennsylvania Perelman School of Medicine Cell and Developmental Biology Microscopy Core maintained these microscopes. Images were analyzed using FIJI (Schindelin et al., 2012) and Volocity Software (PerkinElmer).

Animals

An 18.8-kb region including part of the murine *Trim58* gene locus was retrieved from Bacterial Artificial Chromosome (BAC) bMQ116f01 (Children's Hospital Oakland Research Institute, CHORI) and subsequently cloned into recombineering vectors using

materials obtained from the National Cancer Institute Biological Resources Branch. Artificial loxP sites were inserted to flank exon 3, and Cre recombinase was transiently expressed in E.coli to excise exon 3. This purified construct constituted the *Trim58* exon 3 deletion targeting cassette. The *Trim58* gene was disrupted in Sv129 embryonic stem (ES) cells by transient expression of the targeting cassette and subsequent recombination into the intact genomic locus. ES cells were analyzed via Southern blot for the expected pattern after StuI restriction enzyme digestion. A single ES cell clone was injected into FVB murine blastocysts, and chimeric offspring were crossed with FVB mice. A PCR-based strategy was used to genotype these mice for *Trim58*. The mice used to generate data in this study were subsequently backcrossed at least 4 generations on a C57Bl/6 background for additional analysis. The Children's Hospital of Philadelphia and St. Jude Children's Research Hospital Institutional Animal Care and Use Committees approved all animal protocols.

Peripheral blood and tissue analysis

Peripheral blood was collected in EDTA-coated tubes via retro-orbital sinus or tail vein. Complete blood counts were determined using the Hemavet 950 (Drew Scientific). For phlebotomy studies, 200 μL of blood were collected on days 0 and 2, and fluid volume was replaced with intraperitoneal injection of sterile saline. For peripheral blood analysis of entire blood volume, cardiac puncture was performed under anesthesia. Blood was washed once with PBS, and layered onto a Percoll/NaCl density gradient (1.096-1.058 g/mL) and centrifuged at 250g for 30 min. Purified reticulocytes

populations were validated by flow cytometry using thiazole orange Retic-Count stain (BD).

For phenylhydrazine treatment, 50mg/kg 1-acetyl-2-phenylhydrazine (Sigma) was injected intraperitoneally on two consecutive days. Serial hematocrit levels were determined by standard methods after removing 20 μL of blood from the tail.

Purification of Ter119^{+} erythroid cells

Fetal liver and bone marrow erythroblasts were purified by magnetic sorting using the STEM CELL reagents for PE positive selection with Ter119-PE antibody.

Effect sizes of mouse and human TRIM58 alleles

We used the student *t*-test to calculate the effect sizes of RBC number and MCV in *Trim58*$^{-/-}$ versus *Trim58*$^{+/+}$ mice. We used beta values (β) from GWAS-reported variants associated with the same traits to determine the effect sizes (rs38111444 for RBC number and rs11204538 for MCV)(Kamatani et al., 2010; van der Harst et al., 2012). Beta values (β) were converted to Cohen's D effect sizes using the following formula: Cohen's D = β/sd(β), where sd(β) = sqrt(N/se(β)).

gRNAs and Constructs

gRNA-1: GCTTGTCAAGGCTATTGGTCA

gRNA-2: GTGTCAAGGCTATTGGTCAAG

gRNA sequences were selected using a CRISPR design tool (www.crispr.mit.edu), generated as oligonucleotides, and cloned into plasmids using BbsI. The all-in-one expression plasmid spCas9(BB)-2A-GFP (PX458) was a gift from Feng Zhang (Addgene plasmid #48138). AIO-mCherry vector was cloned by replacing T2A-GFP with IRES-mCherry. The lentiCRISPRv2 was a gift from Feng Zhang (Addgene plasmid #52961). Lentiviral CRISPR-GFP and –mCherry plasmids used in this study were generated by replacing the puromycin resistance gene with IRES-GFP or IRES-mCherry.

Erythroid differentiation of human peripheral blood $CD34^+$ cells

Circulating G-CSF-mobilized human $CD34^+$ cells were obtained from two deidentified healthy donors (Key Biologics, Lifeblood) and enriched by immunomagnetic bead selection using an AutoMACS instrument (Miltenyi Biotec).

$CD34^+$ cells were cultured in a three-phase erythroid differentiation protocol consisting of IMDM (Gibco) supplemented with 2% human AB plasma, 3% human AB serum, 1% penicillin/streptomycin, 3 units/mL heparin, 10 μg/mL insulin, and 3 units/mL erythropoietin (EPO) (Amgen). Phase I (days 1-7) also included 200 μg/mL Holo-Transferrin (Sigma-Aldrich), 10 ng/mL stem cell factor (SCF) (PeproTech, Inc.) and 1 ng/mL IL-3 (PeproTech, Inc.). Phase II (days 8-12) included the same cytokines, except

that IL-3 was withdrawn. During phase III (days 13 and beyond) Holo-Transferrin was increased to 1 mg/mL, and SCF was removed. Erythroid differentiation and maturation were monitored by flow cytometry using anti-CD71-PE (BD Biosciences, clone M-A712), anti-CD235-FITC (BD Biosciences, clone GA-R2), anti-Band3-APC (gift from Xiuli An, NY Blood Center), anti-α_4-integrin-VioBlue (Miltenyi, clone MZ18-24A9).

HUDEP Cell culture

HUDEP clone 2 (HUDEP-2) cells were cultured as previously described. Cells were expanded in StemSpan SFEM (Stem Cell Technologies) supplemented with 1 μM dexamethasone, 1 μg/mL doxycycline, 50 ng/mL human SCF, 3 units/mL EPO, and 1% penicillin/streptomycin. HUDEP-2 cells were differentiated in the Phase III medium used for $CD34^+$ erythroid cultures. These cells tested negative for mycoplasma.

Generation of genomic deletions

$CD34^+$ cells were transduced by centrifuging with lentivirus (multiplicity of infection 40) at 2,800 rpm, 37°C for 90 min with 8 μg/mL polybrene in Phase I erythroid differentiation medium ($CD34^+$) cells or HUDEP-2 expansion medium (for HUDEP-2 cells). The cells were incubated overnight with virus and were switched to fresh medium the following morning. For transient editing, the Amaxa 2b (Lonza; program U-008) was used to electroporate 1-5 million $CD34^+$ or HUDEP-2 cells with DNA plasmid containing gRNA, Cas9, and GFP (2 or 10 μg).

To generate edited HUDEP-2 clones, cells were electroporated with 10 μg DNA plasmid encoding gRNA-1, Cas9, and GFP. Single GFP^+ cells were then sorted into 96-well plates after 24 hours and expanded for 14-21 days. Genomic DNA was isolated using a DNA extraction buffer (100 mM Tris HCl, pH 8.3; 200 mM NaCl; 5 mM EDTA; 1% Triton X-100; 200 μg/mL) proteinase K) followed by incubation at 50°C for 1 hour, then 85°C for 30 minutes.

HbF protein analysis

Erythroid cells derived from $CD34^+$ cells were fixed with 0.05% glutaraldehyde and permeabilized with 0.1% TritonX-100. Cells were stained with anti-HbF-APC antibody (Invitrogen, HBF-1 clone) and analyzed by flow cytometry. HPLC quantification of HbF was performed using a cation-exchange column (Primus Diagnostics).

***In vitro* sickling assay**

Human SCD $CD34^+$ HSPCs (genotype HbSS) were purified from deidentified, discarded whole blood from partial exchange RBC transfusions (considered "not human subject research" by the St. Jude Children's Hospital Institutional Review Board), differentiated into erythroblasts and cultured in 2% oxygen levels between days 14 and 17 of culture (de Vasconcellos et al., 2014). Phase contrast microscopy (CKX41 microscope, Olympus) was performed within 10 minutes of exposure to atmospheric oxygen and analyzed by manual counting of sickled cells, with blinding to sample

genotype. The cell permeable nuclear stain Hoechst 33342 (Sigma) was used to quantify nucleated cells by flow cytometry.

Methylcellulose colony assays

Human $CD34^+$ cells were electroporated with GFP, Cas9 and gRNA expression plasmid using an Amaxa Nucleofector 2b (Lonza), program U-008 and cultured for 24 hours. Live GFP^+ cells were purified by fluorescent activated cell sorting (FACS) for GFP expression and seeded at 300 cells/ml into cytokine-free human methylcellulose (Stemcell) (300 cells/3 cm dish) supplemented with 2 U/mL EPO, 10 ng/mL SCF, 1 ng/mL IL-3, and 1% penicillin/streptomycin. Individual BFU-E colonies were picked after two weeks of culture.

Deep sequencing of genome modifications

DNA was extracted from cells using Blood and Tissue DNA Extraction Kits (Qiagen). PCR amplification using CloneAmp HiFi Premix was performed with primers including Nextera adapters. An additional PCR was performed to individually index each sample, followed by sequencing on a MiSeq platform (Illumina) with 250bp, paired-end reads. The sequence alignment and mutation detection were performed using CLC Genomics Workbench (CLC Bio). *HBG1/2* PCR primers used are as follows:

HBG1-specific Fwd: CGCTGAAACTGTGGTCTTTATGAAAATT

HBG2-specific Fwd: GCACTGAAACTGTTGCTTTATAGGAT

HBG common Rev1 (with *HBG*-specific Fwd): GGCGTCTGGACTAGGAGCTTATTG

HBG-nonspecific Fwd: ATAACCTCAGACGTTCCAGAAGCGAGTGTG

HBG common Rev2 (with *HBG*-nonspecific Fwd):
AGAAGTCCTGGTATCCTCTATGATGGGAG

qPCR detection of 5.2-kb deletion alleles

A primer and probe set was designed to detect amplification of a *HBG1* promoter-specific sequence. TaqMan qPCR was performed on genomic DNA samples from HUDEP-2 and $CD34^+$ cells using Universal TaqMan Mix (Thermo Fisher Scientific) for quantification of triplicates for each sample. ΔΔCt values were calculated based on amplification of RNaseP (Thermo Fisher Scientific) for copy number reference.

5.2kb Fwd: ACGGATAAGTAGATATTGAGGTAAGC

5.2kb Rev: GTCTCTTTCAGTTAGCAGTGG

Taqman probe (FAM): ACTGCGCTGAAACTGTGGTCTTTATGA

Fluorescence *in situ* hybridization (FISH)

A 5.2-kb probe encompassing the intervening region between gRNA-1 cleavage sites in *HBG2* and *HBG1* was generated by PCR amplification and cloned using TA vector (Promega). Nick translation was used to label purified DNA with red-dUTP

(Alexa-Fluor-594, Molecular Probes), and a control *HBB* probe (RP11-1205H24) independently with green-dUTP (Alexa-Fluor-488, Molecular Probes). The probes were hybridized simultaneously with interphase and metaphase cells in 50% formamide, 10% dextran sulfate, and 2× SCC. Metaphase cells were stained with DAPI and scored for signals representing the potentially deleted region (red) and *HBB* (green).

Statistical analyses

Pairwise comparisons were assessed using an unpaired two-tailed Student's *t*-test. Results were considered significant when P - value < 0.05. Linear regression analysis was performed assess potential correlation between BFU-E colony mutation frequencies and γ-globin ratio. Tests were performed and graphed using Prism software (GraphPad).

Chapter 3 The role of Trim58 in erythropoiesis

We initially became interested in studying Trim58 because of its identification by genome-wide association studies (GWAS) as a protein that potentially regulated human erythrocyte traits. GWAS do not distinguish causal variants from blocks of SNPs in linkage disequilibrium, however, so follow-up studies were required to test whether *TRIM58* variation was responsible for changes in RBC formation and function. Moreover, Weiss lab alum Eugene Khandros, M.D., Ph.D., found that Trim58 co-immunoprecipitated with protein aggresomes containing insoluble α-globin. Together, these forward genetic and biochemical screens suggested that Trim58 contributed to erythropoiesis and potentially in the setting of disease. In this work, we identified Trim58 as an erythroid-specific E3 ubiquitin ligase with induced expression during late erythropoiesis. We performed short hairpin (sh) RNA-mediated knockdown studies in primary fetal liver erythroblasts to investigate its role in terminal maturation and enucleation. These studies were performed with Christopher Thom, M.D., Ph.D. and were published in *Developmental Cell* (2014). Figures and Tables can be found at the end of the text.

Additional co-authors who contributed to this work:

Division of Hematology at the Children's Hospital of Philadelphia:
Eugene Khandros, Jenna M Nickas, Olivia Y Zhou, Dolly Prabhu, Yu Yao and Mitchell J Weiss

Department of Physiology and the Pennsylvania Muscle Institute at the University of Pennsylvania Perelman School of Medicine:
Jacob E Lazarus and Erika L Holzbaur

School of Molecular Bioscience at the University of Sydney:
Ana PG Silva and Joel P Mackay

Introduction

Humans produce about 2 million red blood cells each second to replace those lost by senescence (McGrath and Palis, 2008) in a process termed erythropoiesis. Mature red blood cells (erythrocytes) contain high levels of the O_2 carrier hemoglobin and are distensible enough to withstand repeated passages through narrow capillary beds. These properties are acquired during precursor maturation through specialized cell divisions associated with activation of erythroid specific genes (Cheng et al., 2009; Welch et al., 2004), repression of alternate lineage genes (Cheng et al., 2009; Kingsley et al., 2013; Welch et al., 2004), global DNA demethylation (Shearstone et al., 2011), reduced cell volume (Dolznig et al., 1995) and ejection of the nucleus to yield a reticulocyte, which develops further into a mature red blood cell (Liu et al., 2010). The mechanisms that govern these changes in cell structure and morphology are incompletely understood.

The process of nuclear expulsion from erythroblasts (Keerthivasan et al., 2011; Konstantinidis et al., 2012) occurs exclusively in mammals and may represent an evolutionary adaptation to optimize erythrocyte rheology for transport through small capillary beds (Gaehtgens et al., 1981a; Mueller et al., 2008). Erythroblast enucleation is preceded by nuclear condensation and requires histone deacetylation (Ji et al., 2010; Popova et al., 2009) and suppression of the *Myc* proto-oncogene (Jayapal et al., 2010). The condensed nucleus polarizes to one side of the cell via transport mechanisms that require microtubules and phosphoinositide 3-kinase (Wang et al., 2012), followed by Rac GTPase-mediated formation of a contractile actomyosin ring between the incipient

reticulocyte and nucleus (Ji et al., 2008; Konstantinidis et al., 2012). Forces generated by the contractile ring promote further nuclear extrusion (Ji et al., 2008; Wang et al., 2012). Final separation between the reticulocyte and nucleus is facilitated by transport of lipid vesicles to the interface, which facilitates remodeling and resolution of the plasma membrane surrounding both structures (Keerthivasan et al., 2010; 2011). The ejected nucleus with surrounding plasma membrane is termed a pyrenocyte (McGrath and Palis, 2008). *In vivo*, free pyrenocytes are scavenged by macrophages (Soni et al., 2006), while reticulocytes are released into the circulation and undergo further maturation (Gifford et al., 2006). This process shares several features with cytokinesis (Keerthivasan et al., 2010; Konstantinidis et al., 2012). Moreover, all of the proteins currently known to participate in erythroblast enucleation are ubiquitously expressed and most function in mitosis. How erythroblasts co-opt generalized mitotic machinery for the lineage specific process of enucleation is unknown. Presumably, this is mediated in part by the expression of one or more erythroid restricted proteins.

Here, we identify a role for the erythroid protein Trim58 in erythroblast enucleation. Trim58 is a member of the tripartite motif-containing family of proteins, whose members function as E3 ubiquitin ligases broadly in physiology and disease (Napolitano and Meroni, 2012). Genome-wide association studies (GWAS) show that single nucleotide polymorphisms (SNPs) linked to the human *TRIM58* gene associate with variations in the size and/or number of circulating erythrocytes (Kamatani et al., 2010; van der Harst et al., 2012). Here we identify Trim58 as an erythroid restricted

protein that facilitates erythroblast enucleation. Mechanistic studies suggest that Trim58 exerts this function by inducing degradation of the microtubule motor dynein. These findings are the first to identify an erythroid specific protein that participates in erythroblast enucleation, likely by targeting a ubiquitous protein complex that is essential for normal function in virtually all other eukaryotic cells.

Results

Trim58 is induced during late stage erythropoiesis

High-level human *TRIM58* expression is specific to the erythroid lineage (Figures 3.1A and 3.1B) and is strongly induced during the late stages of maturation (Figure 3.1C). *In situ* hybridization analysis of embryonic day 14.5 (E14.5) mouse embryos showed predominant *Trim58* mRNA expression in the fetal liver, an erythropoietic tissue (Figure 3.2A). Semiquantitative real-time PCR showed that Trim58 mRNA is upregulated >100-fold in late stage murine fetal liver erythroid precursors (Figure 3.2B). Chromatin immunoprecipitation-sequencing (ChIP-Seq) of primary erythroblasts demonstrated that the essential erythroid transcription factors Gata1 and SCL/Tal1 bind the *Trim58* locus within the first intron, a common location for erythroid specific enhancers (Figure 3.2C) (Cheng et al., 2009; Pimkin et al., 2014). Thus, the *Trim58* gene is strongly and specifically induced during terminal erythroid maturation, in part via direct activation by key hematopoietic transcription factors.

Trim58 regulates erythroblast enucleation

We used shRNAs to suppress *Trim58* expression during the *in vitro* maturation of primary fetal liver erythroblasts (Zhang et al., 2003). We infected purified E14.5 murine fetal liver erythroid precursors with retroviruses encoding *Trim58*-directed or control shRNAs along with a puromycin resistance cassette and green fluorescent protein (GFP) (Figure 3.3A) (Hemann et al., 2003). Infected cells were cultured for 1-3 days with dexamethasone, stem cell factor (SCF), erythropoietin (Epo) and puromycin to promote immature erythroblast expansion and select for infected cells. The erythroblasts were then switched to medium containing Epo only, which induced maturation to the reticulocyte stage. Four different shRNAs reduced *Trim58* mRNA and protein by 60-90% (Figures 3.3B and 3.3C). During late erythroid maturation, erythroblasts expel their nuclei to become anucleate reticulocytes that are Hoechst-negative by flow cytometry (Figure 3.4A). The kinetics of enucleation were delayed in Trim58-deficient cultures (Figure 3.4B), and enucleation was consistently inhibited at 48 hrs maturation by all four Trim58-directed shRNAs analyzed compared to controls (Figures 3.4C). Histological staining confirmed these findings, showing reduced proportions of reticulocytes in *Trim58*-deficient cultures (Figures 3.4D and 3.4E). *Trim58* suppression also increased the proportions of mature erythroblasts containing two or more nuclei (Figure 3.5).

Several parameters of erythroid maturation were not altered by *Trim58* knockdown, including downregulation of the cell surface marker CD44 (Figure 3.6)

(Chen et al., 2009). *Trim58* knockdown produced only small and inconsistent effects on erythroblast proliferation (Figure 3.7A) and viability (Figure 3.7B). Hemoglobin accumulation (Figure 3.8A) and nuclear condensation (Figure 3.8B) occurred normally after Trim58 knockdown. Overall, these findings demonstrate that Trim58 depletion caused selective defects during late stage erythropoiesis, including reduced enucleation and also increased formation of multinucleated cells.

Trim58 binds the molecular motor dynein

Trim58 is predicted to be an E3 ubiquitin ligase with characteristic functional modules, including a PRY-SPRY (PS) domain that mediates substrate interactions (Figure 3.9A) (James et al., 2007; Woo et al., 2006). We performed pull down studies to identify Trim58 protein-binding partners, including potential degradation substrates. We used the isolated PS domain for these studies because ectopic expression of full-length wild type (WT) Trim58 was toxic to erythroblasts (data not shown). We expressed FLAG epitope-tagged PS domain in the erythroblast cell line G1E (Weiss et al., 1997), which contains no endogenous Trim58 protein, immunoprecipitated (IP) with FLAG antibody and analyzed the recovered proteins by SDS-polyacrylamide gel electrophoresis. This analysis identified five discrete protein bands (Figure 3.9B). Mass spectrometry of these bands identified multiple subunits of the cytoplasmic microtubule motor protein complex dynein (Table 3.1), as well as nuclear pore complex proteins, a Golgi component, and several other proteins (Table 3.2). Dynein regulates nuclear positioning and microtubule

structure within cells (McKenney et al., 2010; Splinter et al., 2010) and it was already known that erythroblast enucleation is microtubule-dependent (Konstantinidis et al., 2012; Wang et al., 2012). Therefore, to investigate further a potential role for Trim58 in erythroblast enucleation, we focused on its interactions with dynein.

We confirmed the Trim58-dynein interaction in G1E cells by FLAG-IP followed by Western blotting for dynein intermediate chain (DIC) and dynein heavy chain (DHC) (Figure 3.9C). Immunoprecipitation of DIC from mouse fetal liver recovered Trim58, demonstrating an interaction between the endogenous proteins (Figure 3.9D). *In vitro* pull down assays showed that recombinant GST-tagged PS interacted with purified holodynein, indicating a direct and specific interaction (Figure 3.10A). Dynein is a multi-subunit complex that includes DHC, DIC, light intermediate chains, and light chains. The DIC mediates dynein interactions with several accessory proteins, mainly via the amino terminus (McKenney et al., 2011; Vallee et al., 2012). To test whether Trim58 interacted with DIC, we expressed hemagglutinin (HA)-tagged segments of the DIC amino terminus in 293T cells, incubated lysates with GST proteins and glutathione-Sepharose beads, and analyzed interacting polypeptides (Figure 3.10B). Western blotting for HA showed that PS interacted with DIC through an interface within the first 73 amino acids. We then used size-exclusion chromatography coupled to multiangle laser light scattering (SEC-MALLS) to confirm this interaction *in vitro* using purified PS and DIC (residues 1-120). Bacterially expressed PS and GST-DIC(1-120) each eluted in gel filtration as a single peak when run individually (Figure 3.10C). MALLS data, which provide a reference-free

estimate of solution molecular weight, showed that PS eluted as a monomer (observed MW = 25.6 kD; expected mass = 23.3 kD), whereas GST-DIC(1-120) ran as a dimer (observed MW = 91.2 kD; expected mass of dimer = 79.6 kD), as a result of the known dimerization of GST (Fabrini et al., 2009). A ~1:1 mixture of the two proteins also eluted predominantly as a single peak with a lower retention time, as expected for the formation of a complex. The molecular weight of this signal calculated from MALLS data (126.2 kD) is consistent with a 2:2 complex (expected mass = 114.4 kD) that represented a 1:1 Trim58-DIC complex that additionally dimerized through the GST moiety. Together, these data show that Trim58 binds dynein directly through physical interactions between its PS domain and the DIC amino-terminus (Figure 3.10D). This region of DIC contains a coiled coil domain that interacts with other dynein regulatory proteins, including the dynactin subunit $p150^{Glued}$ and NudE (Ma et al., 1999; McKenney et al., 2011).

Trim58 promotes dynein degradation

Next, we expressed FLAG/mCherry-tagged full length WT Trim58 in HeLa cells, which do not express endogenous Trim58, and assessed the effects on dynein protein levels. As controls, we expressed vector alone (V) or a "RING-dead" (RΔ) Trim58 containing two missense mutations predicted to abrogate E3 ligase activity (Figure 3.11A) (Zhang et al., 2012). The cultures were treated with puromycin for 3 days to enrich for infected cells and analyzed by Western blotting (Figure 3.11B). Although WT Trim58 slowed cell proliferation, we were able to create viable, stably expressing lines.

Compared to the RΔ mutant, WT Trim58 was poorly expressed, likely due its autoubiquitylation and subsequent proteasomal degradation, which occur with other Trim proteins (Versteeg et al., 2013). Dynein subunits DHC and DIC were nearly absent from cells expressing WT Trim58, but not from cells expressing the RΔ version or the vector alone (Figure 3.11B and data not shown). In order to assess the mechanism through which Trim58 expression causes dynein loss, Trim58-expressing HeLa cells were treated with the proteasome inhibitor MG132. As shown in the right-hand panel of Figure 3.11B, MG132 increased the expression of dynein and WT Trim58. Thus, Trim58 promotes the degradation of itself and of dynein via mechanisms that require E3 ubiquitin ligase activity and proteasomes.

We then investigated whether Trim58 expression elicited cellular phenotypes characteristic of dynein loss. Dynein transports Golgi bodies along microtubules into perinuclear stacks at the microtubule organizing center (MTOC) (Quintyne et al., 1999). Inhibition of dynein by overexpression of CC1, a portion of $p150^{Glued}$ that interferes with the DIC-dynactin interaction, causes Golgi fragmentation and dispersal throughout the cytoplasm (Quintyne et al., 1999). We verified this result in HeLa cells and showed that WT Trim58, but not the RΔ mutant, produced a similar effect (Figure 3.12). Dynein also functions in mitotic spindle checkpoint inactivation to facilitate anaphase onset (Howell et al., 2001). Ectopic expression of WT Trim58 in HeLa cells caused mitotic defects characteristic of dynein inhibition, including prolonged interval between cell rounding and anaphase (Figures 3.13A and 3.13B) and cell death after rounding (Figure 3.13B

right panel) that were significantly worse than cells expressing vector or RΔ Trim58. Thus, ectopic expression of Trim58 in heterologous cells induces phenotypes that are consistent with dynein deficiency.

Trim58 expression in erythroblasts coincides with loss of dynein and enucleation

We investigated whether endogenous Trim58 expression promotes dynein loss during erythropoiesis. In murine erythroblast cultures undergoing semi-synchronous maturation, the induction of Trim58 protein at 24 hours correlated with loss of dynein subunits (DIC and DHC, Figure 3.14A) and the onset of enucleation (Figure 3.14C, black bars). The absence of dynein in late stages of erythroid maturation is consistent with proteomic studies of murine (Pasini et al., 2008) and human (Goodman et al., 2007) erythrocytes. In contrast, Trim58-deficient cells retained dynein protein aberrantly (Figures 3.14B and 3.15) and enucleation was delayed (Figures 3.14C gray bars). Dynein expression still declined in Trim58 knockdown erythroblasts, but with a ~12 hr delay. The eventual degradation of most dynein by 44 hours may be due to incomplete suppression of Trim58 and/or alternate protein degradation mechanisms. Overall, our findings demonstrate that expression of Trim58 destabilizes dynein and promotes enucleation. These findings are consistent with a model in which Trim58-mediated degradation of dynein and erythroblast enucleation are mechanistically linked.

We then used immunofluorescence to examine relationships between microtubule structure and nuclear positioning during enucleation. Erythroblast nuclei are surrounded by a network or "cage" of microtubules, similar to what occurs in most cells (Figure 3.16) (Tsai et al., 2010; Wilson and Holzbaur, 2012). During enucleation, the nucleus moves away from the MTOC (Figure 3.20 and (Konstantinidis et al., 2012; Wang et al., 2012)), opposite to what would occur with minus end-directed dynein transport. Later, the microtubule cage partially collapses, becoming detached from the cell cortex and nucleus as the latter is extruded (Figures 3.16ii-iv). Dynein stabilizes microtubules and tethers them to the cell cortex (Hendricks et al., 2012). Thus, events observed during enucleation, including directional nuclear movement and microtubule cage collapse (Figure 3.17), may be facilitated by dynein loss.

Chapter commentary

In this chapter, we performed a carefully controlled characterization of the effects of Trim58-targeting shRNA effects in a primary cell culture model of erythroid differentiation. These data indicated that Trim58-deficiency or the shRNAs themselves delayed erythroid enucleation, likely by disrupting the kinetics of dynein degradation. We also performed imaging-flow cytometry studies to dissect the specific nuclear morphology changes and movements (condensation, polarization, and extrusion) that were inhibited by (what we now know are caused by) Trim58-directed shRNAs

themselves. To test the effects of complete *Trim58* loss on erythropoiesis and substrate degradation, the next step was generating and analyzing a *Trim58* knockout mouse model.

Figures

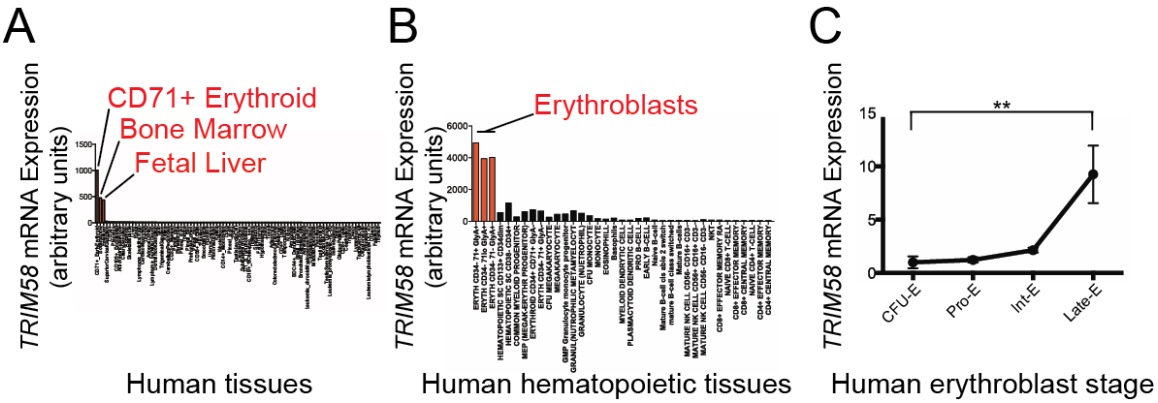


Figure 3.1. Human Trim58 expression is restricted to late stage erythroblasts.

Human *TRIM58* mRNA expression levels were measured by microarray analysis in (A) a panel of human tissue types (Wu et al., 2009), (B) a panel of human hematopoietic tissues (Novershtern et al., 2011), and (C) cultured primary human erythroblasts from peripheral blood buffy coat mononucleocytes (Merryweather-Clarke et al., 2011). The results in panel C represent mean ± SD for 3 biological replicates normalized to the average expression level in CFU-E cells, which were assigned a value of 1. **$p<0.01$. CFU-E, erythroid colony-forming units; Pro-E, proerythroblasts; Int-E, intermediate (basophilic) erythroblasts; Late-E, late (orthochromatic) erythroblasts.

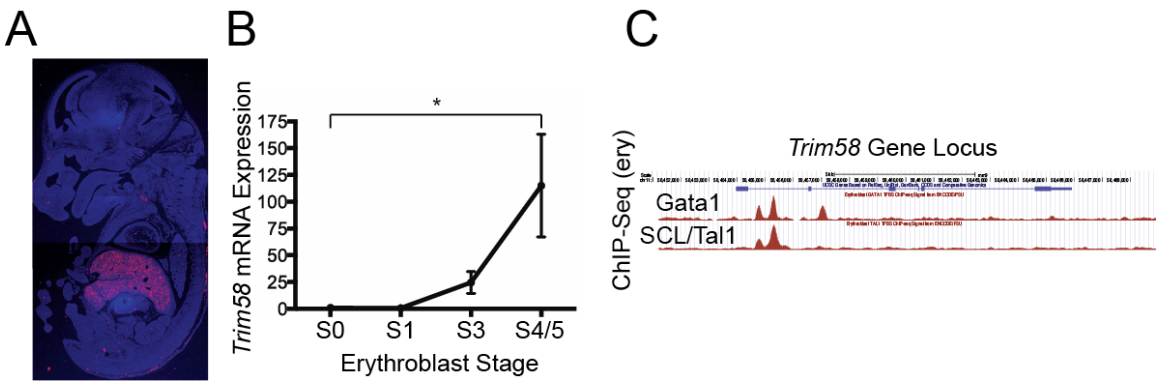


Figure 3.2. Murine Trim58 is expressed during late stage erythropoiesis.

(A) RNA *in situ* hybridization for *Trim58* mRNA in a day 14.5 mouse embryo showing strong expression (red) in fetal liver, the site of definitive erythropoiesis. (B) Primary murine fetal liver erythroblasts were FACS-purified using a previously described gating strategy (Pop et al., 2010) and *Trim58* mRNA was analyzed by semiquantitative real-time PCR. The y-axis shows relative mRNA expression, normalized to S0 cells, which were assigned a value of 1. The results represent mean ± SEM for 3 biological replicates. *$p<0.05$. (C) Chromatin immunoprecipitation-sequencing (ChIP-Seq) analysis of transcription factor binding to the *Trim58* locus in primary murine erythroblasts (data from Pimkin *et al*, in revision). The blue line depicts the *Trim58* gene, with exons shown as rectangles. Transcription factor binding sites are indicated in red.

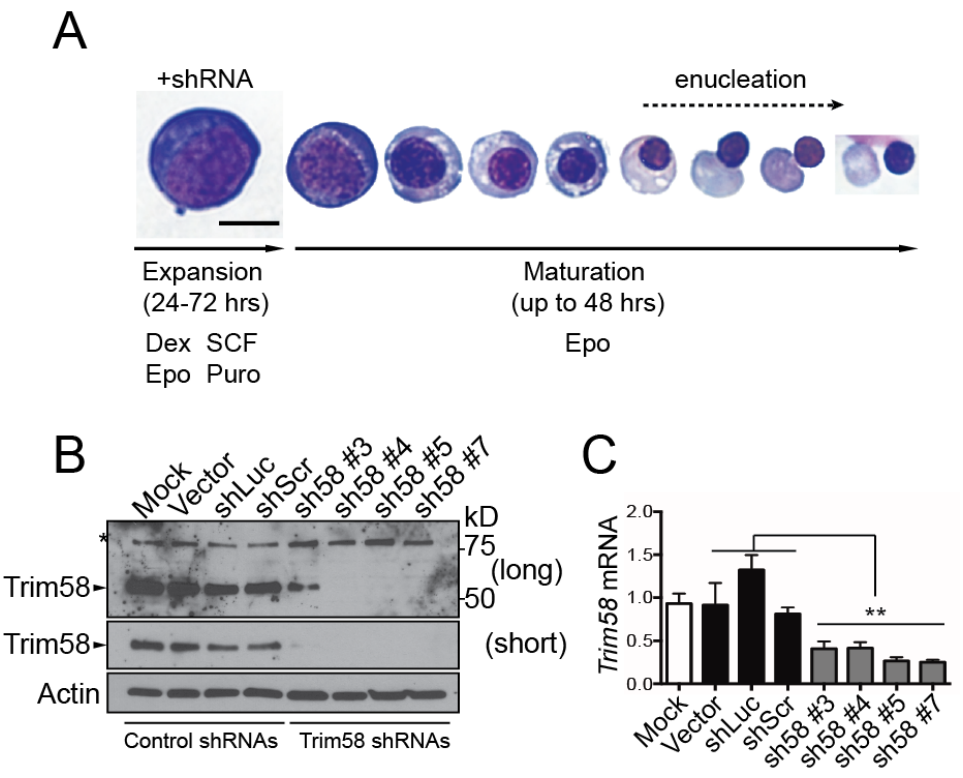


Figure 3.3. Trim58-directed shRNAs reduce expression in murine erythroblast cultures.

(A) Embryonic day 14.5 murine fetal liver erythroid precursors were purified, infected with retroviruses encoding *Trim58* or control shRNAs, and cultured for 24-72 hrs in expansion medium with puromycin (Puro) to select for infected cells. The cells were then switched to maturation medium, which facilitates development to the reticulocyte stage over ~48 hrs. Epo, erythropoietin; SCF, stem cell factor; Dex, dexamethasone. Scale bar, 10mm. (B) Western blot for Trim58 in erythroblasts expressing Trim58-directed or control shRNAs at 48 hrs maturation. Luc, luciferase; Scr, scrambled. The asterisk (*) represents a nonspecific band. "Long" and "short" exposures are from the same blot. (C) Suppression of *Trim58* mRNA expression by four shRNAs. *Trim58* mRNA levels were

measured at 48 hrs maturation by semiquantitative real time PCR and normalized to the average expression value for *Gapdh*, *Actin*, and *Hprt* mRNA levels. The y-axis represents mRNA expression level compared to the average value for control-treated samples (Vector, shLuc, shScr), which was set at 1. Results represent mean ± SD for 4 biological replicates.

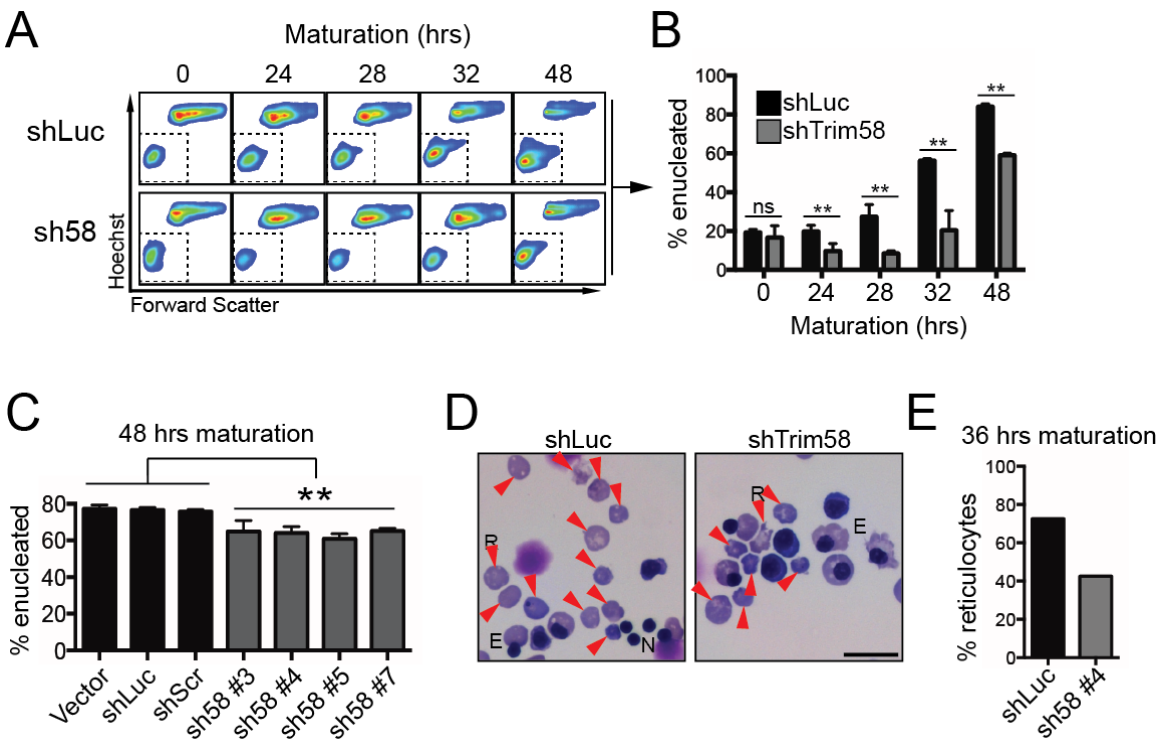


Figure 3.4. Trim58 knockdown inhibits murine erythroblast enucleation.

FACS analysis for enucleation of control (shLuc) and Trim58-deficient (shTrim58 #4) erythroid cultures at the indicated time points. During maturation, erythroblasts become smaller, indicated by decreased forward scatter, and ultimately enucleate to become Hoechst-negative reticulocytes (boxed regions). (B) Summary of percent (%) enucleation at the indicated time points in cultures treated with control (shLuc) or Trim58-directed (#4) shRNAs. Results represent mean ± SD for 3 biological replicates. (C) Four Trim58-directed shRNAs inhibit enucleation. Following 48 hrs maturation, erythroblasts treated with the indicated shRNAs were analyzed by FACS for enucleation. (D) Representative histologies of control (shLuc) and Trim58-deficient (shTrim58 #4) erythroid cultures

after 36 hrs maturation. Red arrows denote anucleate reticulocytes. Scale bar, 20 mm. R, reticulocyte; E, erythroblast; N, extruded nucleus. (E) Summary of cell counts from panel D showing percent (%) reticulocytes. Six hundred cells were counted on each slide.

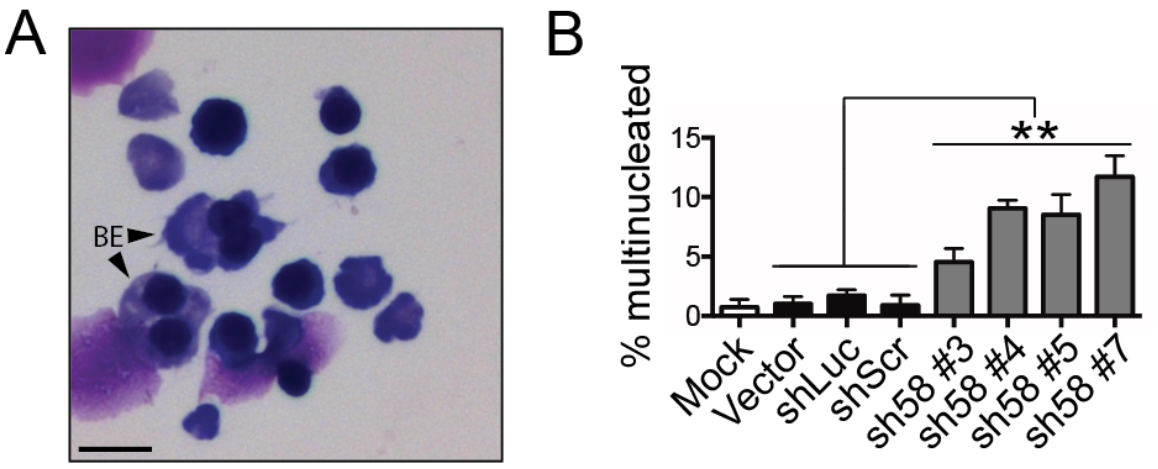


Figure 3.5. Trim58 knockdown results in multinucleated erythroblasts.

Murine fetal liver erythroblasts were infected with control retrovirus(es) (empty Vector; shLuc, Luciferase; and/or shScr, Scrambled), *Trim58*-directed shRNA(s) (sh58 #3, #4, #5, and/or #7) and/or mock-treated. Infected cells were expanded with puromycin for 72 hrs and then induced to mature at time 0. (A) At 48 hrs maturation, erythroblasts treated with shTrim58 #4 were cytocentrifuged onto a glass slide, stained with May Grünwald and Giemsa reagents, and visualized by light microscopy. Arrows denote binucleate erythroblasts (BE). Scale bar, 10 mm. (B) The percentage of cells containing 2 or more nuclei was quantified by visual inspection. Results represent mean ± SD for 3 independent experiments with >200 cells counted per treatment group per experiment.

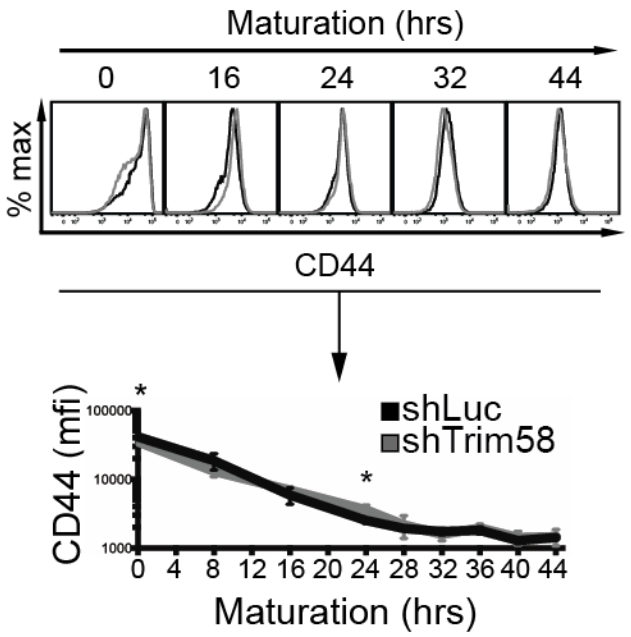


Figure 3.6. Trim58 knockdown does not affect CD44 downregulation kinetics during erythroid maturation.

Erythroblasts expressing control (shLuc, black) or Trim58 (sh58 #4, gray) shRNAs were analyzed serially by flow cytometry for CD44, a cell surface marker that progressively decreases during erythroid maturation (Chen et al., 2009). Representative histogram plots of FACS data are shown. Below the histograms are line graphs depicting CD44 mean fluorescence intensity (mfi) over time (mean ± SD for 4 biological replicates). % max, percent (%) of maximum cell count.

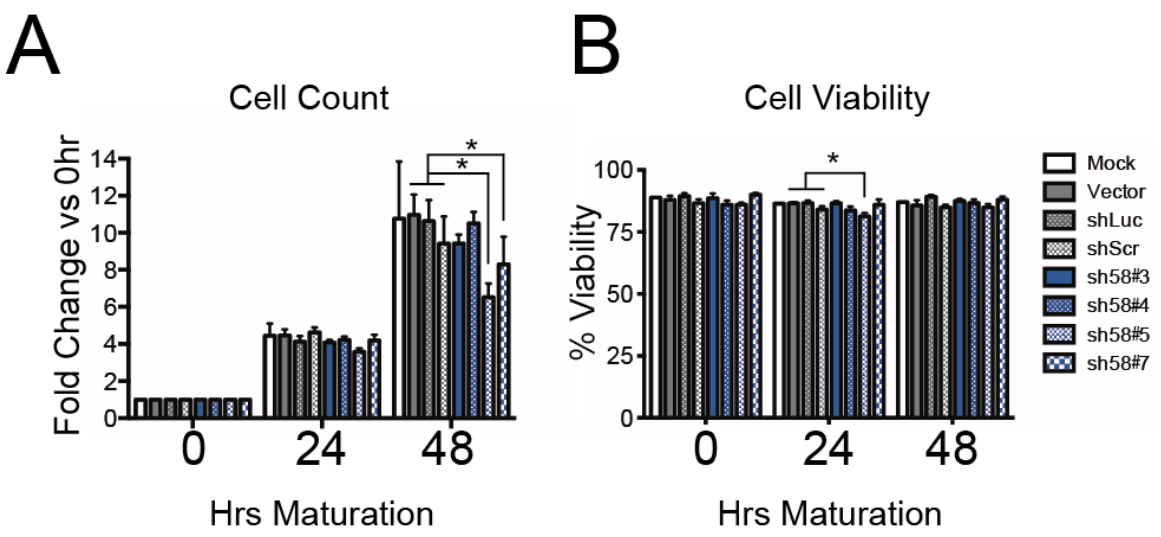


Figure 3.7. Trim58 knockdown does not affect cell proliferation or viability.

Murine fetal liver erythroblasts were infected with control retrovirus(es) (empty Vector; shLuc, Luciferase; and/or shScr, Scrambled), *Trim58*-directed shRNA(s) (sh58 #3, #4, #5, and/or #7) and/or mock-treated. Infected cells were expanded with puromycin for 72 hrs and then induced to mature at time 0. (A) Infected erythroblast proliferation was quantified at 0, 24, and 48 hrs maturation. Cell numbers were normalized to the cell count at 0 hrs. Results represent mean ± SD for 4 biological replicates. (B) Cell viability was measured by fluorescent amino-reactive (LiveDead) staining at 0, 24, and 48 hrs maturation. Results represent mean ± SD for 4 biological replicates. A key depicting the color of each treatment group is shown to the right.

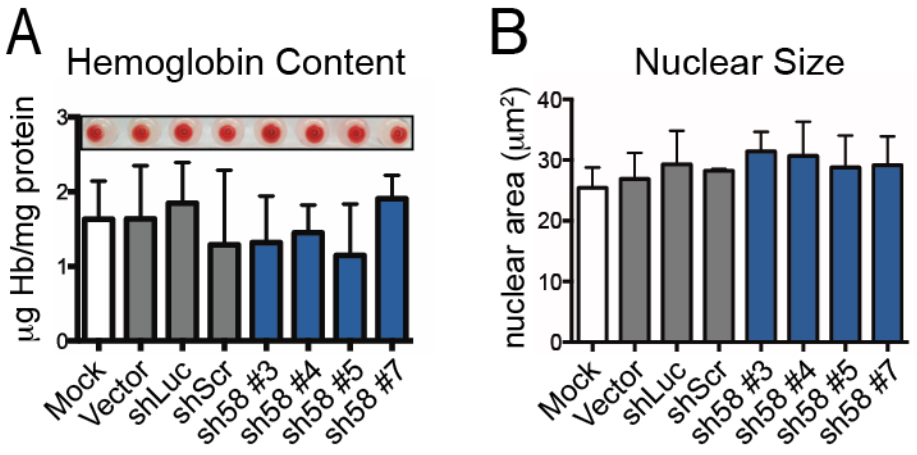


Figure 3.8. Trim58 knockdown does not affect Hb content or nuclear condensation.

Murine fetal liver erythroblasts were infected with control retrovirus(es) (empty Vector; shLuc, Luciferase; and/or shScr, Scrambled), *Trim58*-directed shRNA(s) (sh58 #3, #4, #5, and/or #7) and/or mock-treated. Infected cells were expanded with puromycin for 72 hrs and matured for 48 hours. (A) Hemoglobin content measured at 48 hrs maturation, normalized to total cellular protein. Results represent mean ± SD for 3 independent experiments. The top panel shows the red color of cell pellets. (B) Nuclear size quantification at 48 hrs maturation. Erythroblasts treated with the indicated shRNAs were centrifuged onto a slide and stained with May Grünwald and Giemsa reagents. The area encompassed by each nucleus was quantified by microscopy using the Analyze Particles feature in FIJI (Schindelin et al., 2012). Results represent mean ± SD for 3 independent experiments with >50 cells counted per treatment group per experiment. *$p<0.05$, **$p<0.01$.

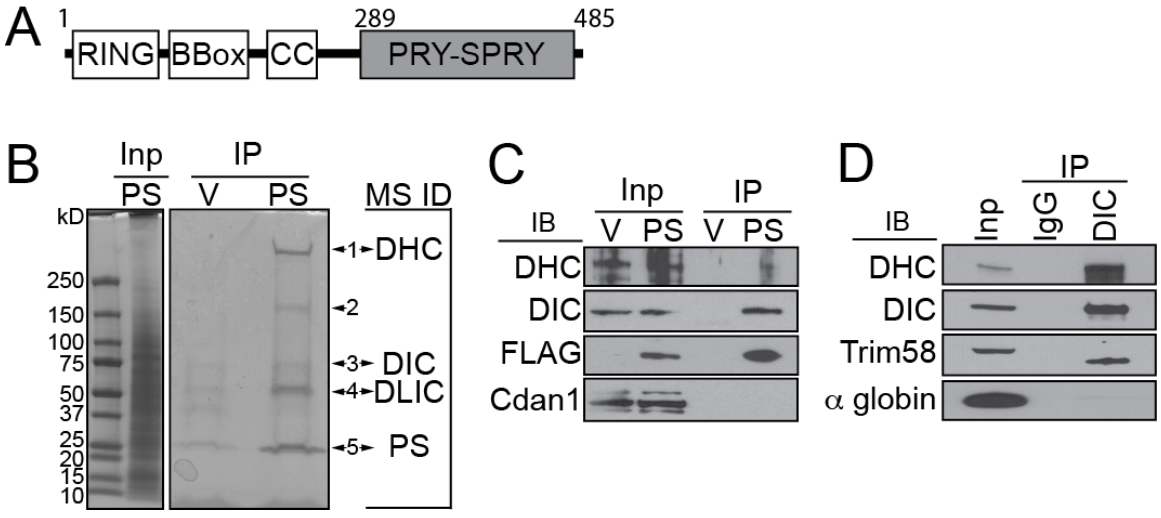


Figure 3.9. The Trim58 PRY-SPRY domain binds the molecular motor dynein.

(A) Domain structures of Trim58 showing the "tripartite" RING-BBox-Coiled Coil (CC) motif and the immunoglobulin-like PRY-SPRY domain. Amino acid numbers are indicated. In some TRIM proteins, the RING domain recruits E2 conjugases carrying activated Ubiquitin and the PRY-SPRY domain binds substrates. (B) FLAG-tagged Trim58 PRY-SPRY domain (PS), or vector (V), were stably expressed in G1E proerythroblast cells. Lysates were immunoprecipitated (IP) with FLAG antibody, fractionated by SDS-polyacrylamide gel electrophoresis, and stained with Coomassie Blue Silver. Numbered arrows denote the regions of the gel from which bands were excised for mass spectrometric (MS) analysis. Bands from both V and PS IPs were analyzed. MS identified peptides from dynein heavy chain (DHC), dynein intermediate chain (DIC), dynein light intermediate chains 1/2 (DLIC), and Trim58 PRY-SPRY (PS) predominantly in Bands 1, 3, 4, and 5, respectively. These proteins were only identified

in bands from the PS IP lane. See Tables 3.1 and 3.2 for a full listing of identified proteins. Inp, input (1%). (C) Lysates of G1E cells expressing FLAG-PS or vector (V) were immunoprecipitated with FLAG antibody and analyzed by Western blotting (immunoblotting, IB) for DHC, DIC, FLAG and Codanin1 (Cdan1, negative control). Inp, input (5%). (D) Embryonic day 14.5 murine fetal liver erythroblasts were lysed, immunoprecipitated with anti-DIC antibody or IgG control and analyzed by Western blotting for the indicated proteins. a globin was included as a negative control. Inp, input (0.5%).

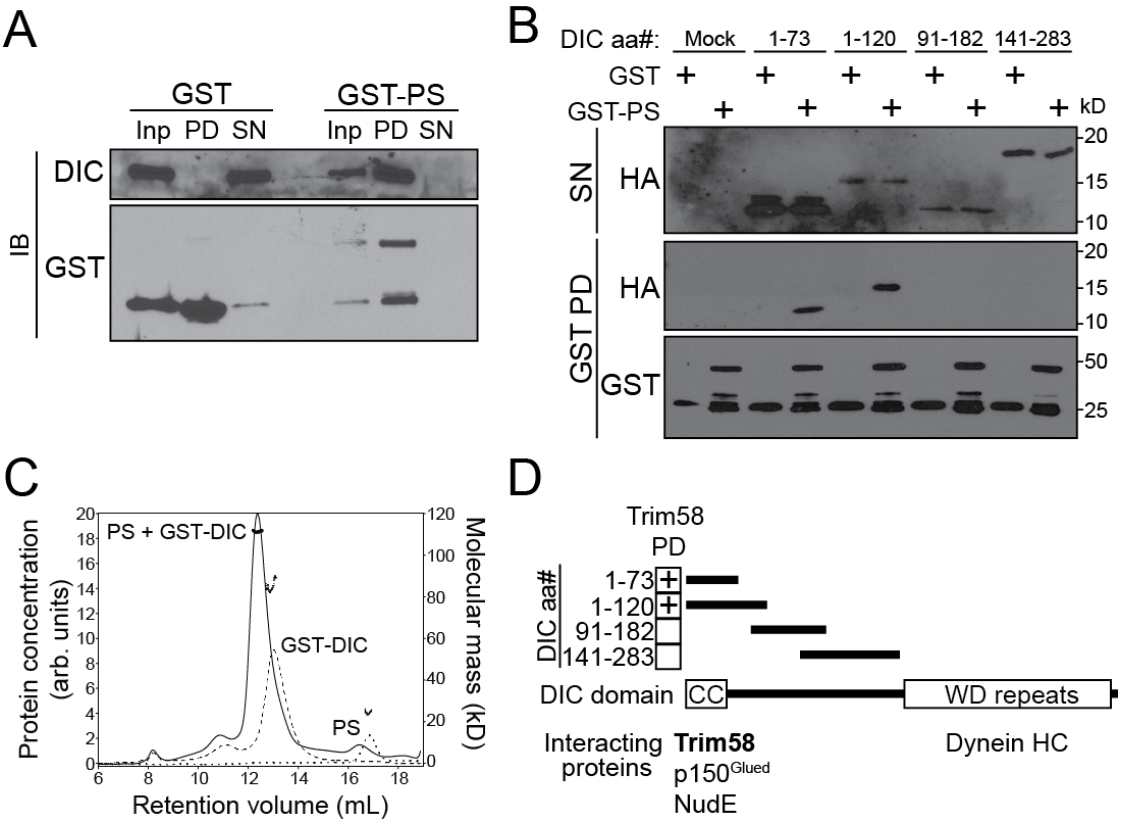


Figure 3.10. Trim58 binds directly to the DIC amino terminus.

(A) Recombinant purified GST-Trim58 PS domain (GST-PS) or GST alone were incubated with purified bovine holodynein complex and glutathione-Sepharose beads. Bound proteins were analyzed by Western blotting. Equal percentages of total input (Inp), pull down (PD) and supernatant (SN) samples were loaded. (B) 293T cells were transfected with expression plasmids encoding hemagglutinin (HA) fused to the indicated DIC amino acids. After 24 hrs, cells were lysed and incubated with GST-PS or GST and glutathione-Sepharose beads. Bound proteins were analyzed by anti-HA Western blotting. PD, pull down. SN, supernatant (1%). (C) SEC-MALLS data for PS, GST-

DIC(1-120) and a 1:1 mixture. Protein concentration was measured on an inline refractive index detector. Light scattering data converted to molecular weight are shown above each chromatography trace and relate to the right-hand y-axis. The observed molecular weights are consistent with a 1:1 interaction between the two proteins augmented by the dimerization of the GST tag appended to DIC(1-120). (D) DIC domain structure showing the amino-terminal coiled coil (CC) motif, which binds Trim58. This region of DIC also binds dynactin/p150^{Glued} and NudE, which are previously characterized regulators of dynein function (McKenney et al., 2011). The DIC carboxy-terminal domain contains WD repeats that interact with DHC. Amino acid numbers are indicated (aa#).

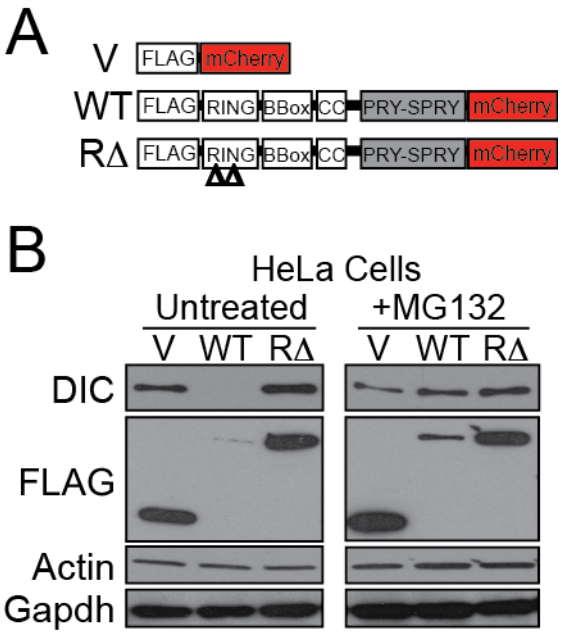


Figure 3.11. Trim58 facilitates dynein degradation.

(A) Schematic diagram showing full length Trim58 constructs with amino-FLAG and carboxyl-mCherry tags. The RD mutant contains two missense mutations in the RING domain that abrogate ubiquitin ligase activity. (B) Western blot analysis of Trim58-expressing HeLa cells. Cells were infected with retrovirus encoding FLAG-Trim58 (WT or RD) or vector (V), selected in puromycin for 3 days, then analyzed by Western blotting before or after treatment with the proteasome inhibitor MG132 (10 mM) for 4 hrs at 37 ºC.

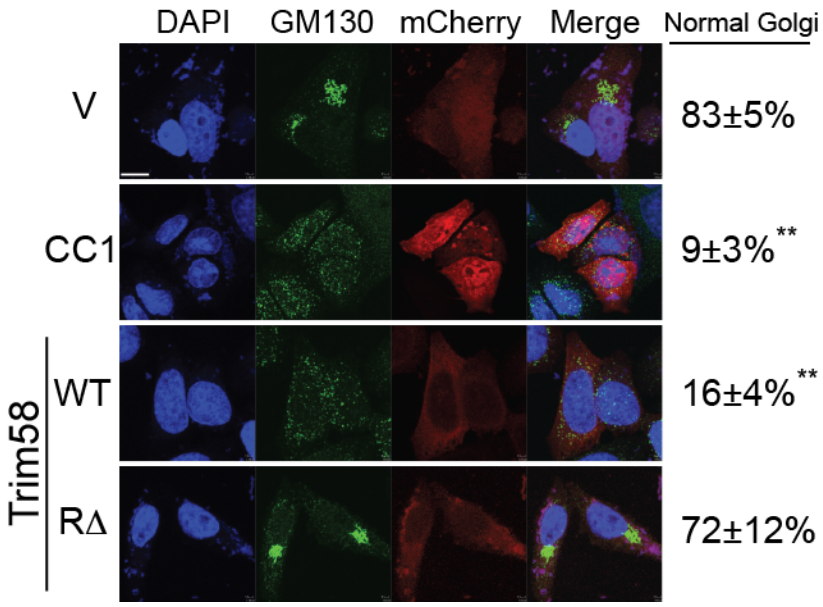


Figure 3.12. Trim58 expression causes Golgi spreading, a marker of dynein dysfunction.

Assessment of Golgi structure in HeLa cells expressing Trim58. HeLa cells were transfected with expression plasmids encoding Trim58-mCherry or CC1-mCherry, a dynein inhibitor (Quintyne et al., 1999). After 36 hrs, the cells were fixed, stained for the Golgi matrix protein GM130 and DNA (DAPI), and visualized by confocal microscopy. The percentage (%) of mCherry (mCh) positive cells that displayed normal punctate perinuclear Golgi body distribution is shown at right. The results represent mean ± SD for 3 independent experiments, with >100 cells counted per experiment. **$p<0.01$ vs. Vector control.

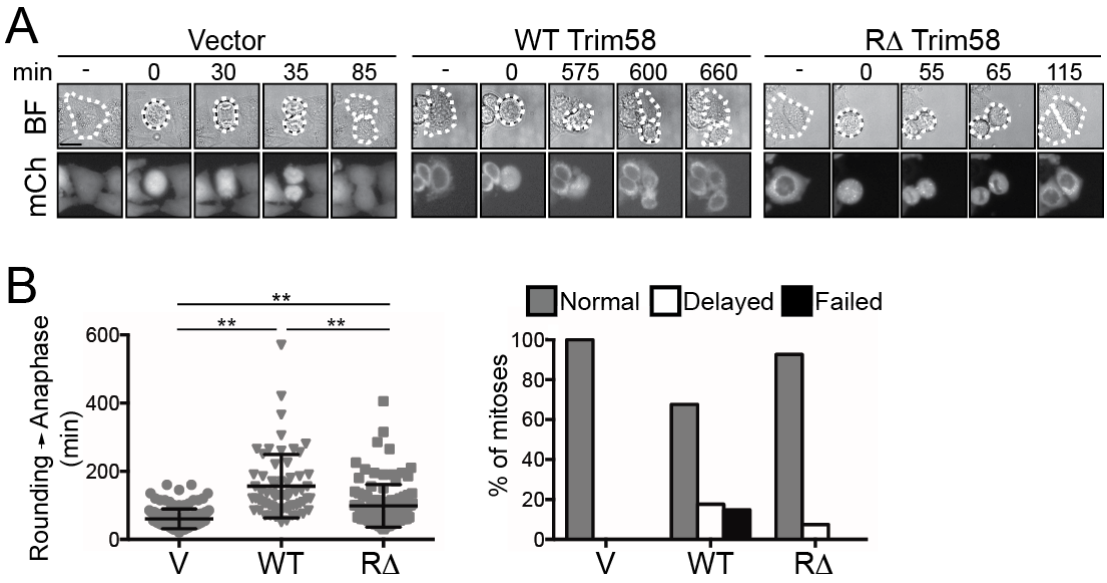


Figure 3.13. Trim58 expression perturbs mitotic progression.

(A) Mitotic progression in HeLa cells expressing Trim58-mCherry constructs was analyzed by time-lapse microscopy. Representative images show mitotic progression in outlined cells. Cells expressing WT Trim58 exhibit delayed progression from cell rounding (~metaphase) to anaphase, compared to cells expressing vector or RD Trim58. BF, Brightfield; mCh, mCherry. Scale bar, 16mm. (B) Quantitative analysis of mitosis in Trim58-expressing HeLa cells. The graph on the left shows time elapsed between cell rounding and anaphase. The results represent mean ± SD for all observed mitotic cells. **$p<0.01$. The graph on the right shows the percentages of cells with delayed (>200 min) or failed mitosis manifested as cell death between rounding and anaphase. (V, n=134 normal mitoses; WT, n=50 normal, 13 delayed, 11 failed mitoses; RD, n=100 normal, 7 delayed mitoses).

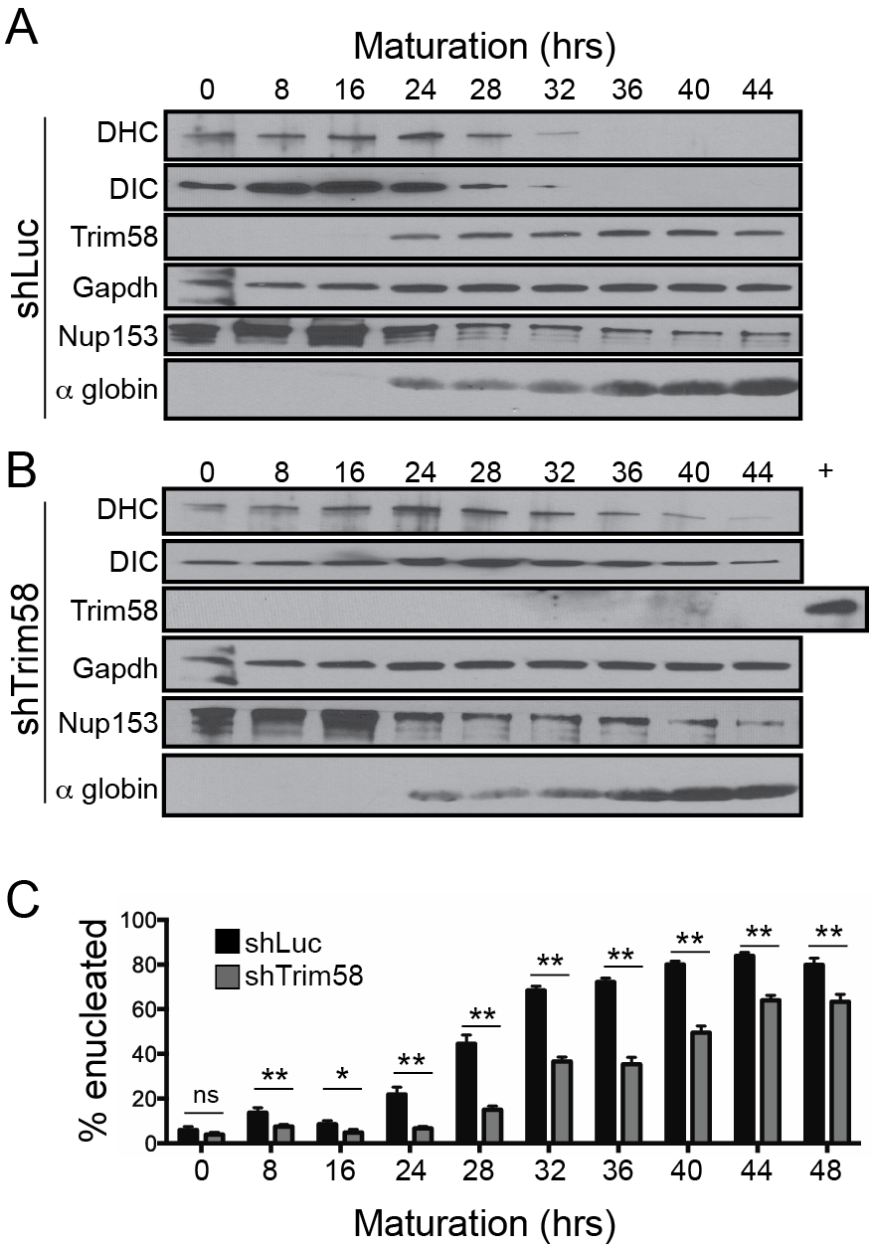


Figure 3.14. Trim58 expression correlates with loss of dynein and enucleation during erythroid maturation.

Fetal liver erythroblasts were infected with retrovirus encoding shRNA against (A) luciferase (shLuc) or (B) Trim58 (shRNA #4), cultured in expansion medium for 72 hrs, and shifted to maturation medium at time 0. Whole cell lysates were prepared at the indicated time points and analyzed by Western blotting. A control sample from 44 hrs maturation was run in the lane marked "+" as a positive control for Trim58. (C) Rate of enucleation in cells from panels A and B, determined as shown in Figure 3.4A. The results represent mean ± SD for 4 biological replicates. *$p<0.05$; **$p<0.01$; ns, not significant.

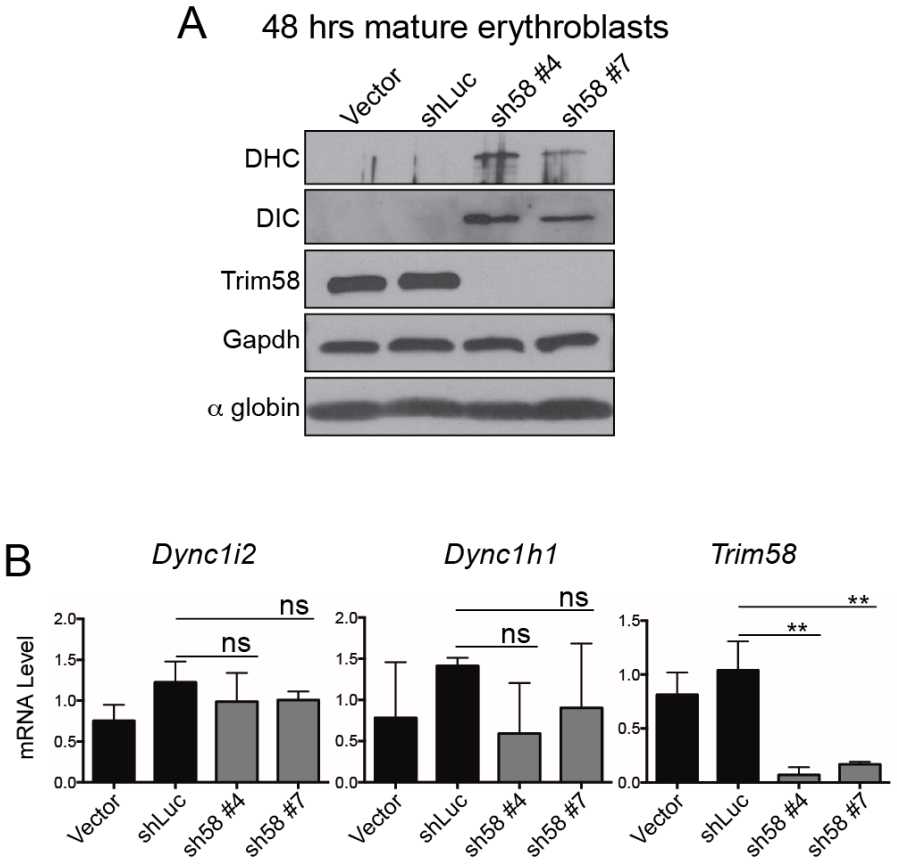


Figure 3.15. Trim58 deficiency causes aberrant dynein protein retention in late stage erythroblasts.

Erythroblast cultures expressing control constructs (Vector, shLuc) or Trim58-directed shRNAs (sh58 #4, sh58 #7) were expanded in puromycin for 72 hrs, then induced to mature for 48 hrs and collected for analysis. (A) Whole cell lysates were analyzed by Western blotting for the indicated proteins. DHC, dynein heavy chain; DIC, dynein intermediate chain. (B) Whole cell lysate cDNA was analyzed by semiquantatitive real time PCR for dynein intermediate chain (*Dync1i2*), dynein heavy chain (*Dync1h1*) and *Trim58*. The y-axis represents mRNA expression level compared to the average value for

control-treated samples (Vector, shLuc), which was set at 1. Results represent mean ± SD for 4 biological replicates. **p<0.01; ns, not significant.

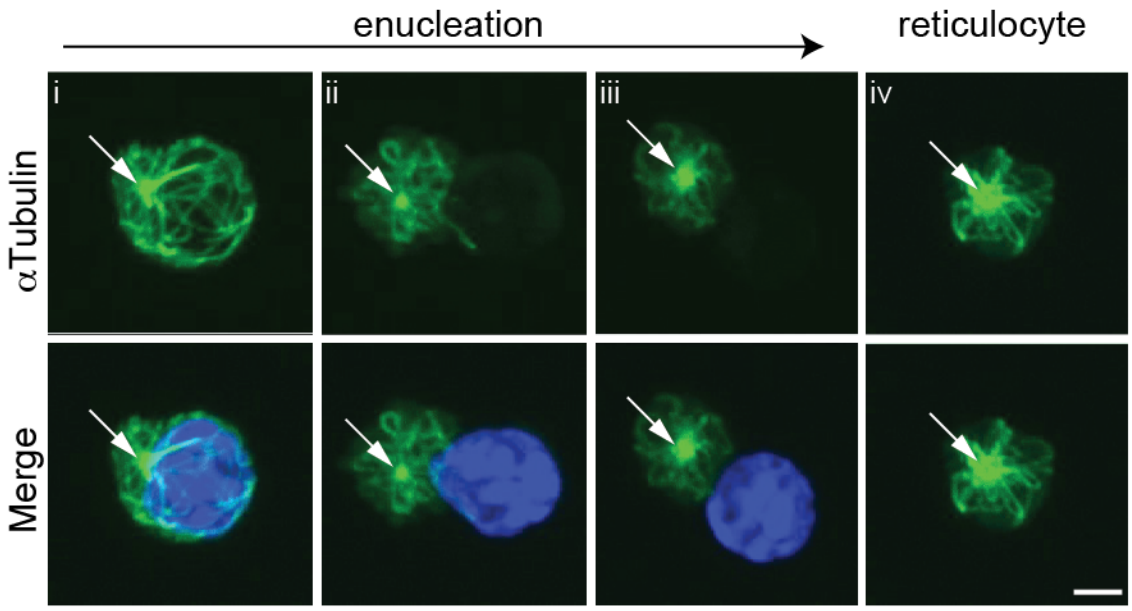


Figure 3.16. Directional nuclear movement during erythroblast enucleation.

Primary murine fetal liver erythroblasts were fixed, stained for microtubules (α tubulin, green) and DNA (DAPI, blue), and imaged by deconvolution fluorescent microscopy. Arrow indicates the microtubule organizing center (MTOC). Scale bar, 2 μm.

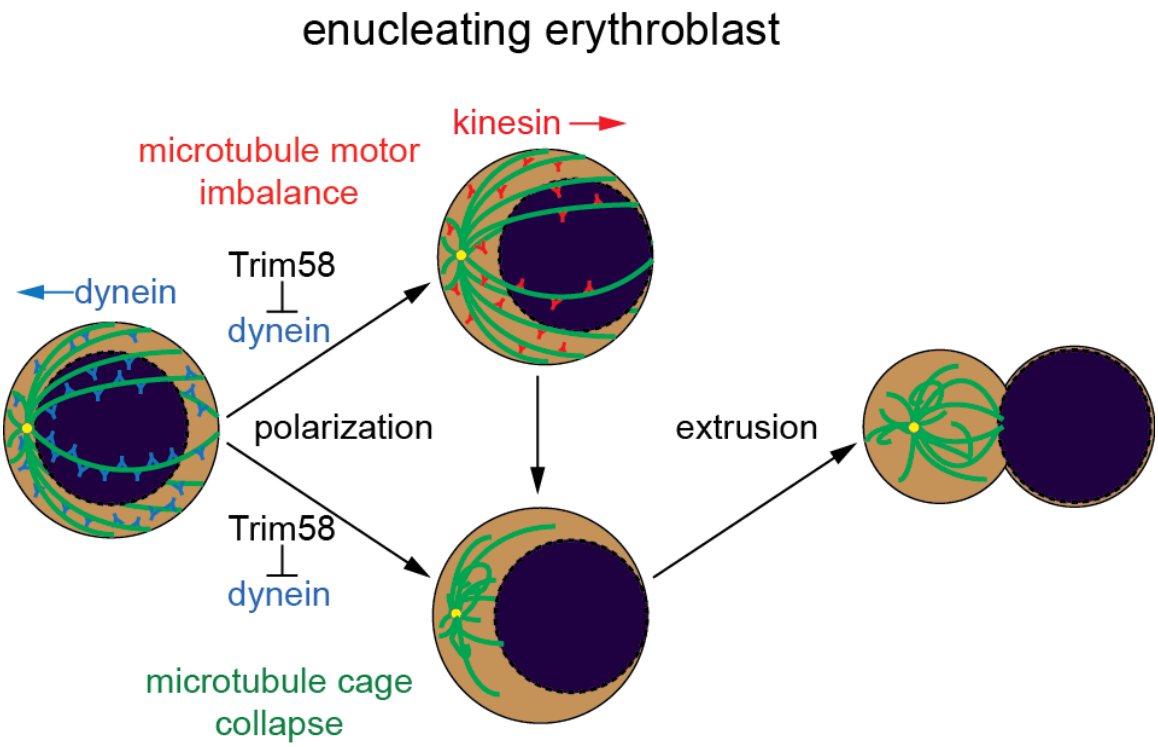


Figure 3.17. Model for the actions of Trim58 during erythroblast enucleation.

In early maturation, the nucleus resides within a cage of microtubules (green) and dynein (blue) keeps the nucleus in close proximity to the MTOC (yellow). Induction of Trim58 causes dynein degradation, which may promote nuclear polarization through multiple mechanisms. A microtubule motor imbalance may allow unopposed kinesin molecular motors to polarize the nucleus within a microtubule cage (top) and/or microtubule cage collapse could free the nucleus to polarize. Following polarization and microtubule disassembly, nuclear extrusion occurs.

Tables

Protein Name	Accession	MW	Band	Unweighted Spectra	Unique Peptides	Percent Coverage
Cytoplasmic dynein 1						
Heavy chain 1	IPI00119876	532 kD	1	189	95	19
Intermediate chain 2	IPI00131086	68 kD	3	16	7	14
Light intermediate chain 1	IPI00153421	57 kD	3,4	38	12	27
Light intermediate chain 2	IPI00420806	54 kD	4	12	7	16
Trim58 PRY-SPRY (Bait)						
Trim58	IPI00353647	55 kD	4,5	91	14	26

Table 3.1. Dynein subunits identified by mass spectrometric analysis of proteins that coimmunoprecipitated with the Trim58 PRY-SPRY domain in erythroid cells.

Chapter 4 The role of Trim58 *in vivo*

In this chapter, we performed experiments to follow up *in vitro* studies implicating Trim58 as a necessary component of erythroblast enucleation. In the previous chapter, we showed that Trim58 stimulated ubiquitination and proteasomal degradation of the dynein molecular motor complex. Expression of *Trim58*-directed short hairpin RNAs (shRNAs) in cultured mouse fetal liver erythroblasts caused abnormal accumulation of dynein and delayed erythroblast enucleation. To examine the functions of *Trim58 in vivo,* we disrupted the gene by removing exon 3, which encodes an essential coiled-coil motif. Figures and Tables can be found at the end of the text.

Additional co-authors who contributed to this work:
Division of Hematology at the Children's Hospital of Philadelphia:
Christopher Thom, Jenna M Nickas, and Olivia Y Zhou

Department of Hematology at St. Jude Children's Research Hospital:
Yu Yao and Mitchell J. Weiss

Department of Computational Biology at St. Jude Children's Research Hospital:
Evadnie Rampersaud, Lance Palmer

Department of Physiology and the Pennsylvania Muscle Institute at the University of Pennsylvania Perelman School of Medicine:
Erika L Holzbaur

Chapter Summary

Trim58 is an E3 ubiquitin ligase expressed at high levels in late stage erythroblasts. Previous work showed that Trim58 stimulated ubiquitination and proteasomal degradation of the dynein molecular motor complex. Expression of *Trim58*-directed short hairpin RNAs (shRNAs) in cultured mouse fetal liver erythroblasts caused abnormal accumulation of dynein and delayed terminal maturation by inhibiting extrusion of the nucleus (enucleation). To examine the functions of *Trim58 in vivo,* we disrupted the gene by removing exon 3, which encodes an essential coiled-coil motif. Erythroblasts, reticulocytes and RBCs of homozygous targeted mice lacked Trim58 protein and contained abnormally elevated dynein levels. The mutant mice showed mild alterations in red blood cell size and number, consistent with genome wide association studies (GWAS) linking the human *TRIM58* gene to these traits. However, *Trim58*$^{-/-}$ mice had no anemia at steady state and recovered normally after phenylhydrazine-induced hemolysis. Moreover, in contrast to results obtained using RNA interference in cultured wild type erythroblasts, *Trim58*$^{-/-}$ erythroblasts enucleated normally in the same assays. These studies show that Trim58 eliminates dynein expression during erythroid maturation *in vivo*. However, this process is largely dispensable for erythropoiesis.

Introduction

Mature red blood cells are void of organelles and contain mainly hemoglobin, which makes up approximately 95% of total cellular protein (Roux-Dalvai et al., 2008). To achieve this specialized state, streamlining of the proteome occurs during erythroid maturation via multiple mechanisms including transcriptional and post transcriptional alterations in gene expression, and selective protein elimination, in part through autophagy and the ubiquitin proteasome system (UPS). The UPS is a multi-enzyme cascade that catalyzes the formation of polyubiquitin chains onto proteins at specific lysine residues, thereby targeting them for degradation. Generalized inhibition of the UPS by drugs or RNA interference impairs erythropoiesis (Chen et al., 2002; Egan et al., 2015; James et al., 2007). Indeed, anemia is a recognized toxicity of bortezimib, an FDA-approved proteasome inhibitor used for treating multiple myeloma (San Miguel et al., 2006).

The UPS facilitates selective elimination of proteins, including those that are irreversibly damaged, cytotoxic, or unnecessary (Khandros and Weiss, 2010). This pathway degrades target proteins through a multistep mechanism. An E1 enzyme activates ubiquitin and transfers it to an E2 ubiquitin-conjugating enzyme. E3 ubiquitin ligases act as scaffolds to bind the E2-ubiquitin enzyme and substrate, thereby promoting its polyubiquitination and subsequent recognition by the proteasome, a multi-subunit organelle that includes numerous proteases. Hundreds of different E3 ligases, each recognizing a limited repertoire of substrates, confer specificity to the UPS.

Transcriptome and proteomic studies showed that UPS components, including numerous E3 ligases, are upregulated during late stages of erythroid maturation (An et al., 2014; Egan et al., 2015; Pasini et al., 2006). However, relatively few of these E3 ligases have been studied with respect to substrate recognition and functions during RBC formation (Khandros and Weiss, 2010; Wefes et al., 1995).

We recently characterized Trim58, an E3 ubiquitin ligase that is strongly induced during late stage erythropoiesis (Thom et al., 2014). In human GWAS, *TRIM58*-linked nucleotide polymorphisms segregate with altered RBC size and number, implicating a role in erythropoiesis (Ganesh et al., 2009; van der Harst et al., 2012). *Trim58* is a member of the Tripartite Motif-containing (Trim) superfamily including more than 65 proteins, many of which function as E3 ubiquitin ligases with important roles in cell physiology (Ozato et al., 2008). The N-termini of Trim proteins contain sequential RING (Really Interesting New Gene), B box, and coiled-coil domains, while various substrate binding domains are present at the C-terminus (James et al., 2007). Previously, we showed that the C-terminal PRY-SPRY domain of Trim58 binds the dynein molecular motor complex to facilitate its ubiquitination and degradation *in vitro* and in cultured cells (Thom et al., 2014). Dynein is an essential multi-subunit protein that transports cellular cargo, including proteins, vesicles, and organelles, along microtubules toward the minus-end. Expression of *Trim58*-targeting short hairpin (sh) RNAs in cultured fetal liver erythroblasts prevented dynein elimination during terminal maturation and delayed enucleation that occurs during normal RBC formation. Thus, we proposed that Trim58-

mediated degradation of dynein facilitates erythroid maturation by promoting enucleation.

To examine further the role of Trim58 in erythropoiesis, we disrupted the corresponding gene in mice. Removal of exon 3, which encodes an essential coiled-coil domain, largely eliminated both ubiquitin ligase activity towards dynein and Trim58 expression in erythroblasts. Nucleated erythroblasts, reticulocytes and mature red blood cells from *Trim58*$^{-/-}$ mice contained abnormally elevated levels of dynein. The mutant mice also exhibited altered RBC number and size (mean corpuscular volume, MCV), consistent with GWAS, which linked the human *TRIM58* gene to these traits. However, *Trim58*$^{-/-}$ mice had no anemia at baseline and recovered normally after phenylhydrazine-induced hemolysis. Moreover, terminal maturation of cultured *Trim58*$^{-/-}$ fetal liver and adult erythroblasts, including enucleation, was normal. Furthermore, we showed that *Trim58* shRNAs inhibit enucleation of cultured mouse erythroblasts through sequence-independent off-target effects that vary according to genetic strain. Collectively, our studies demonstrate that Trim58 promotes degradation of dynein *in vivo*, but is largely dispensable for erythropoiesis and thrombopoiesis.

Results

Disruption of the *Trim58* gene.

The first intron of *Trim58* contains a predicted erythroid enhancer identified by a DNase I hypersensitive site and binding of the transcription factors Gata1 and Tal1 (Thom et al., 2014). Our targeting strategy preserved this region, as intergenic enhancers can potentially regulate distant genes (Harmston and Lenhard, 2013). We removed *Trim58* exon 3 (Figures 4.1A and 4.2A), which encodes a coiled-coil domain required for ubiquitin ligase activity of other Trim family members (Sanchez et al., 2014). Mutant mice, referred to as $Trim58^{-/-}$ for reasons discussed below, were born at normal Mendelian ratio, exhibited no obvious morphological abnormalities and survived to adulthood. We verified correct targeting of the *Trim58* locus by Southern blotting (Figure 4.11B), PCR of genomic DNA (Figure 4.1C) and quantitative reverse transcription PCR (RT-PCR) of E14.5 fetal liver cells (Figure 4.2A-B). Homozygous and heterozygous mutant erythroblasts expressed an aberrant *Trim58* mRNA that included exons 1 and 2, but lacked distal sequences (Figures 2A-B and 4.3A). Thus, the mutant gene is probably transcribed at normal rate, but largely processed abnormally with loss of sequences that are distal to deleted exon 3 (Figure 4.2B, right).

Excision of *Trim58* exon 3 creates an in-frame deletion of 77 amino acids. In principle, this could generate a mature RNA encoding an internally deleted form of Trim58 protein that lacks only the coiled-coil region, but contains the RING, B-box and PRY-SPRY domains (Figure 4.3B). This *Trim58* transcript, termed ΔCC, was expressed

in homozygous mutant erythroblasts (Figure 4.3A), albeit at very low levels (Figure 4.2B, right panel). ΔCC or wild type (WT) Trim58 proteins were not visualized in Western blot studies of *Trim58*$^{-/-}$ fetal livers (Figure 4.2C) or circulating reticulocytes from adult mutant animals (Figure 4.10C, below). Nonetheless, we assessed potential functions of the Trim58 ΔCC protein by transducing NIH3T3 cells with retroviruses encoding FLAG/mCherry-tagged versions of either WT Trim58, Trim58 ΔCC, or a "RING-dead" (RΔ) Trim58 containing two missense mutations that eliminate E3 ubiquitin ligase activity (Figure 4.3B-C). In contrast to Trim58$^{-/-}$ erythroblasts, Trim58 ΔCC was stably expressed in NIH3T3 cells. However, unlike WT Trim58, ectopically expressed ΔCC Trim58 failed to eliminate dynein (Figure 4.3C). Overall the mutant protein is expressed in heterologous cells, although it does not degrade dynein, in agreement with studies demonstrating that the CC domain is necessary for ubiquitin ligase activity (Sanchez et al., 2014). However, *Trim58* ΔCC mRNA and protein are expressed minimally in *Trim58*$^{-/-}$ erythroblasts. Thus, we conclude that ablation of exon 3 eliminated endogenous erythroid *Trim58* activity.

Normal hematopoiesis in Trim58-/- mice with a mixed genetic background.

Initially, we analyzed *Trim58*$^{-/-}$ mice with a mixed genetic background (Sv129, FVB, and C57Bl/6). No major differences in erythroid parameters occurred in mutant mice compared to WT littermates, although we observed trends toward decreased RBC

number and increased mean corpuscular volume (MCV) (Table). Fetal liver, bone marrow and splenic erythroid precursors were normal in number and no maturation defects were evident by flow cytometry analysis of cell surface markers Ter119 and CD71 (Figures 4.4 and 4.5). Adult *Trim58*$^{-/-}$ mice exhibited normal recovery after anemia induced by phenylhydrazine (Figure 4.5C).

Trim58$^{-/-}$ erythroblasts undergo normal enucleation in vitro.

Previously, we showed that multiple shRNAs targeting different regions of *Trim58* mRNA inhibited enucleation of cultured WT fetal liver erythroblasts (Thom et al., 2014). We performed similar studies on E14.5 fetal liver erythroblasts from the progeny of intercrossed *Trim58*$^{+/-}$ mice. Erythroblasts were expanded for 1-3 days by culturing in dexamethasone, stem cell factor, and erythropoietin (Epo), followed by removal of the first two components to induce terminal maturation. We monitored erythroblast enucleation by flow cytometry of cells stained with the erythroid antigen Ter119 and Hoechst 33342, a cell permeable nuclear dye (Figure 4.6A). Surprisingly, Trim58$^{-/-}$ fetal liver erythroblasts enucleated normally after 24 and 48 hours of maturation (Figure 4.6B), contrasting with our previous results obtained after expression of Trim58 shRNAs in E14.5 erythroblasts (strain CD1) (reference (Thom et al., 2014) and Figure 4.6C). This apparent discrepancy could be due to strain-specific effects of either *Trim58* deficiency or *Trim58* shRNAs. Two independent *Trim58* shRNAs had no effect on the enucleation of mixed strain (Sv129, FVB, and C57Bl/6) *Trim58*$^{-/-}$ or *Trim58*$^{+/+}$

erythroblasts (Figure 4.7A), despite effective reduction of *Trim58* mRNA in mixed strain WT cells (Figure 4.7B).

Trim58$^{-/-}$ mice with C57Bl/6 background have mild hematological abnormalities.

To investigate further potential strain effects of *Trim58* deficiency, we bred *Trim58*$^{-/-}$ alleles onto a pure C57Bl/6 background. Similar to what we observed on the mixed genetic background, C57Bl/6 *Trim58*$^{-/-}$ mice were not anemic (Table 4.1) and RBC morphologies on blood smears were normal (not shown). *Trim58*$^{-/-}$ C57Bl/6 mice exhibited mild but significantly decreased RBC number and increased MCV (Table 4.1), similar to the trends observed in *Trim58*$^{-/-}$ mixed background mice. Of note, GWAS studies link the human *TRIM58* gene to variations in RBC number, MCV (Gieger et al., 2011; Kamatani et al., 2010; van der Harst et al., 2012). Human GWAS studies also link *TRIM58* allelic variation to altered platelet count; this parameter trended downward in *Trim58*$^{-/-}$ mice, but was not significant ($p=0.09$ and $p=0.2$ for mixed genetic background and C57/Bl6 strains, respectively).

Trim58$^{-/-}$ mice with C57Bl/6 background reveal strain-specific shRNA off-targeting.

Remarkably, two different *Trim58* shRNAs reduced enucleation of C57Bl/6 erythroblasts by 25-35%, regardless of *Trim58* genotype ($^{+/+}$ or $^{-/-}$) (Figure 4.8A). As expected, the shRNAs effectively reduced full-length *Trim58* mRNA in C57Bl/6

erythroblasts (Figure 4.8B). To summarize, *Trim58* gene disruption in mixed or C57Bl/6 backgrounds produced mild alterations in RBC number and MCV, with no effects on *in vitro* enucleation of cultured erythroblasts. In contrast, two shRNAs targeting non-overlapping segments of *Trim58* mRNA inhibited enucleation of cultured CD1 and C57Bl/6 erythroblasts, but had no effects on erythroblasts with mixed genetic background. Thus, the effects of *Trim58* shRNAs on erythroblast enucleation are genetic strain-specific and occur through mechanisms that are independent of *Trim58* suppression.

Increased dynein in Trim58$^{-/-}$ erythroblasts.

To investigate whether Trim58 promotes dynein degradation *in vivo*, we compared its expression in erythroid cells of WT and *Trim58*$^{-/-}$ mice (mixed strain, Sv129/FVB/C57Bl/6). We phlebotomized animals to induce semi-synchronous erythropoiesis (Figure 4.9A) and collected bone marrow and circulating RBCs at various stages of maturation. *Trim58*$^{+/+}$ and *Trim58*$^{-/-}$ erythroblasts were purified from the bone marrow by Ter119 immunomagnetic bead selection; cells of both genotypes were of similar maturation stage, based on Ter119 and CD71 expression (Figure 4.9B). Although dynein is expressed in most cells, levels decrease during erythroid maturation (Thom et al., 2014), as evidenced in the current study by Western blotting for dynein intermediate chain (IC) in WT Ter119^{+} bone marrow erythroblasts (Figure 4.9C). In contrast, *Trim58*$^{-/-}$

erythroblasts expressed dynein IC at a level similar to that of unfractionated bone marrow.

Next, we used a Percoll gradient to purify circulating reticulocytes and mature RBCs (Blanc et al., 2009) (Figure 4.10A). Coomassie blue silver stain showed no obvious differences in the protein content of WT vs. Trim58$^{-/-}$ reticulocytes (Figure 4.10B). Trim58 protein was not detected by Western blotting in homozygous mutant reticulocytes (Figure 4.10C). However, *Trim58*$^{-/-}$ reticulocytes (Figure 4.10C) and RBCs (Figure 4.10D) expressed abnormally elevated levels of dynein IC. Together, these data are consistent with previous findings that Trim58 ubiquitinates dynein IC to facilitate degradation of the dynein complex (Thom et al., 2014). The current data demonstrate that Trim58 promotes dynein degradation during erythropoiesis *in vivo*, although accumulation of dynein upon loss of Trim58 has minimal consequences on RBC formation at baseline or during recovery of induced hemolytic anemia, and no effect on enucleation of cultured erythroblasts.

Chapter commentary

Here, we discovered an unusual phenotype caused by shRNAs, found to be independent of *Trim58* mRNA suppression. This is discussed further in Chapter 6. These studies emphasize the importance of follow-up studies and the caution necessary in interpreting results using a single experimental approach (i.e. shRNA-induced

knockdown in fetal liver erythroblasts). Overall, we felt it was important to gain a deeper understanding of these effects in such a commonly used model system of erythroid differentiation and maturation. This commitment drove us to test the effects of shRNAs on different strains of mice and subsequently backcross the *Trim58*$^{-/-}$ allele onto a penetrant background.

Figures

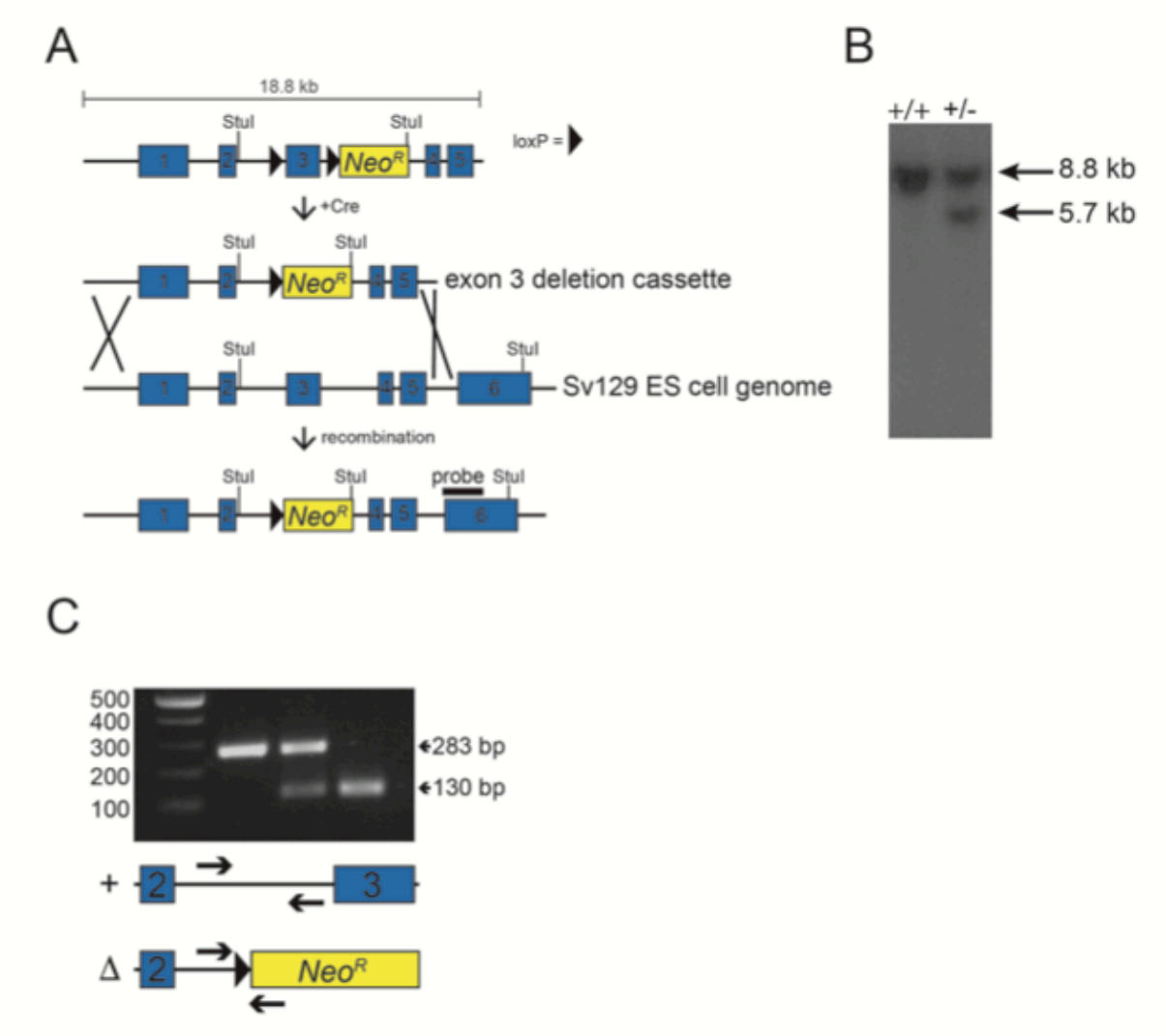


Figure 4.1. Generation of Trim58$^{-/-}$ mice.

(A) Design of Trim58$^{-/-}$ mice. LoxP sites were inserted around exon 3, and Cre recombinase expression mediated excision in the ES cells. (B) Southern blot of targeted ES cell clones with predicted StuI pattern. (C) Validation of PCR amplification used for genotyping genomic DNA.

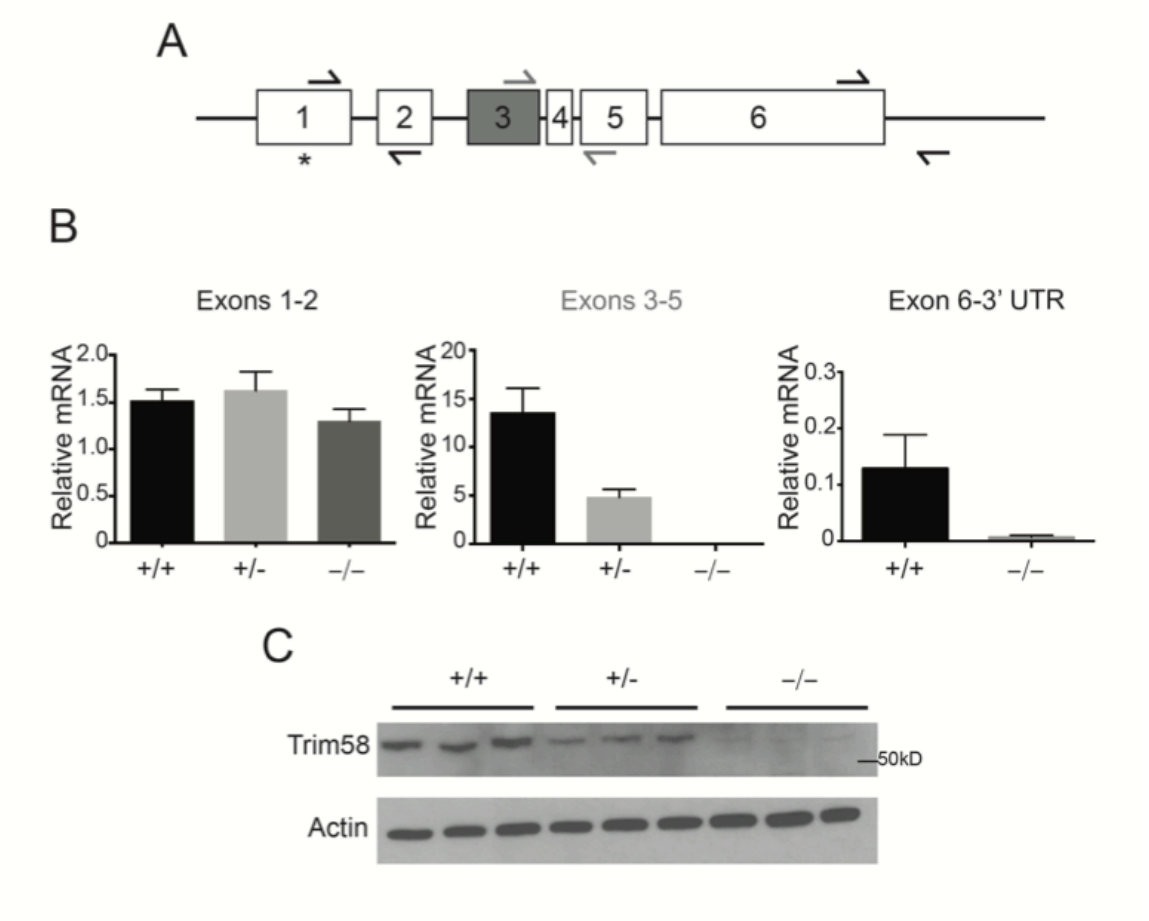


Figure 4.2. Disrupted *Trim58* expression in *Trim58*$^{-/-}$ mice.

(A) *Trim58* gene structure (not drawn to scale). The asterisk denotes the epitope for an anti-peptide antibody (amino acids 7-29). Half arrows indicate primer pairs used for qRT-PCR analysis of mRNA expression (see panel B). Exon 3, which was deleted in the gene targeted mice, and the corresponding PCR primers are shown in grey. (B) Quantitative RT-PCR analysis of E14.5 fetal liver mRNA from wild type ($^{+/+}$), heterozygous ($^{+/-}$), and *Trim58*-null ($^{-/-}$) embryos using the primer pairs shown in panel A. The results represent

mean ± standard deviation for 3 biological replicates. (C) Western blot of E14.5 fetal liver using the anti-peptide antibody described in panel A. β-actin was used as a loading control. Arrow denotes the predicted molecular weight (50 kD) of protein with deleted exon 3.

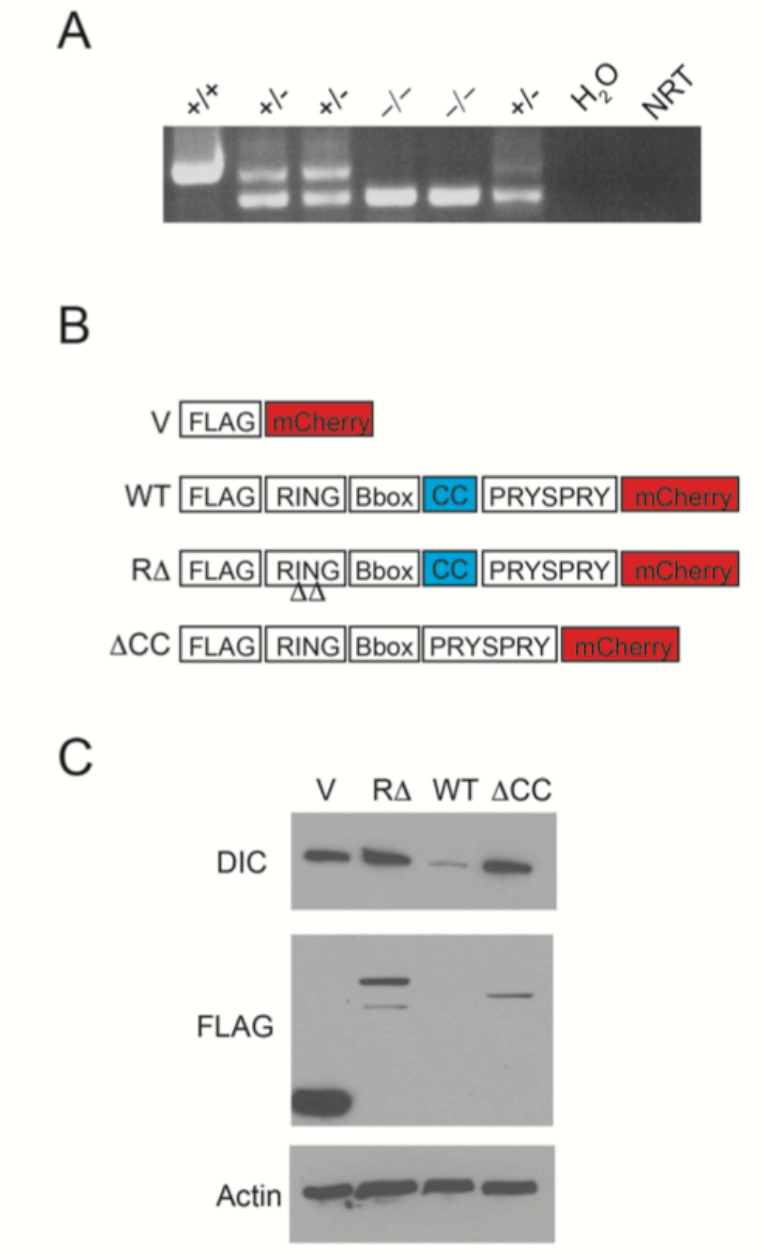


Figure 4.3. *Trim58*$^{-/-}$ mice express an inactive Trim58 allele, ΔCC.

(A) Non-quantitative reverse transcription-PCR of E14.5 fetal liver cells to detect variant *Trim58* mRNA. Higher band represents full-length *Trim58* expression in WT, and the

lower band detected in heterozygotes and ΔCC represents truncated mRNA excluding exon 3. (B) Schematic of *Trim58* expression constructs with N-terminal FLAG and C-terminal mCherry tags. The RΔ mutant contains two missense mutations in the RING domain that abolish E3 ubiquitin ligase activity. The ΔCC cDNA was cloned from *Trim58*$^{-/-}$ fetal liver tissue. (C) Western blot analysis of Trim58-expressing NIH3T3 cells. Cells were infected with retrovirus expressing FLAG-*Trim58* (WT, RΔ, or ΔCC) or empty vector (V), selected in puromycin for 3 days, then analyzed by Western blot for dynein IC, FLAG, and actin for loading control.

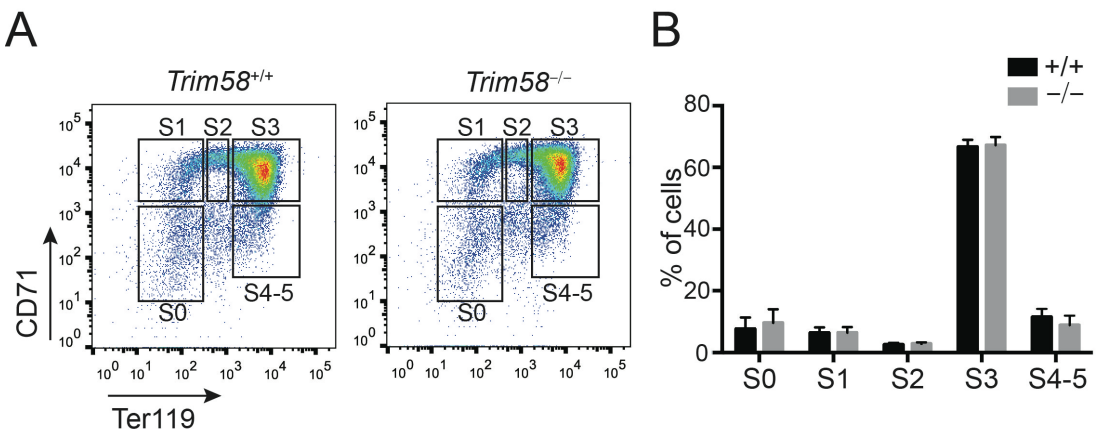


Figure 4.4. Normal erythropoiesis in *Trim58*$^{-/-}$ embryos with mixed genetic background (Sv129, FVB, and C57Bl/6).

(A) Flow cytometry analysis of E14.5 fetal liver cells for CD71 and Ter119 expression. The data shown are representative from two experiments examining two different litters. (B) Quantitation of plots for each population shown in panel C, represented as mean ± SD. n = 7 for each group, from three E14.5 pregnancies. Means were ns by student's t-test.

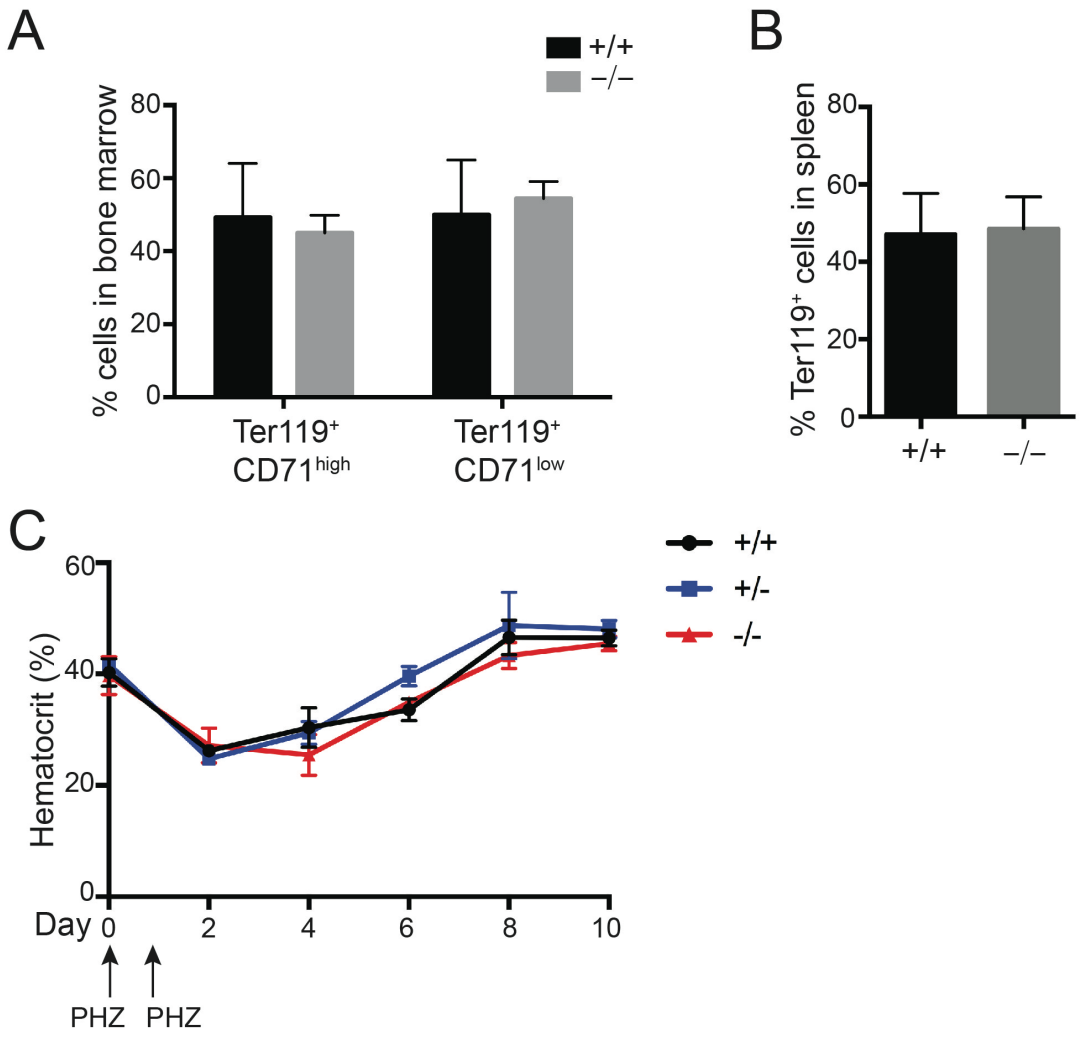


Figure 4.5. Normal adult erythropoiesis in E14.5 *Trim58*$^{-/-}$ mice.

(A) Quantitation of Ter119^{+}, CD71^{high} and Ter119^{+}, CD71^{low} erythroid cells in bone marrow isolated from mice (age 8-12 weeks), evaluated by flow cytometry. Results are shown as mean ± SD (n = 3). Means were ns by student's t-test. (B) Quantitation of Ter119^{+} erythroid cells in spleens isolated from mice (age 8-12 weeks). Results are shown as mean ± SD (n = 3). Means were ns by student's t-test. (C) Recovery of

Trim58$^{+/+}$, *Trim58*$^{+/-}$ and *Trim58*$^{-/-}$ mice after phenylhydrazine (PHZ)-induced hemolysis. Arrows indicate treatments with PHZ (60 mg/kg). Serial hematocrit measurements were performed on peripheral blood. $n = 4$ for each group; each time point was ns by student's *t*-test.

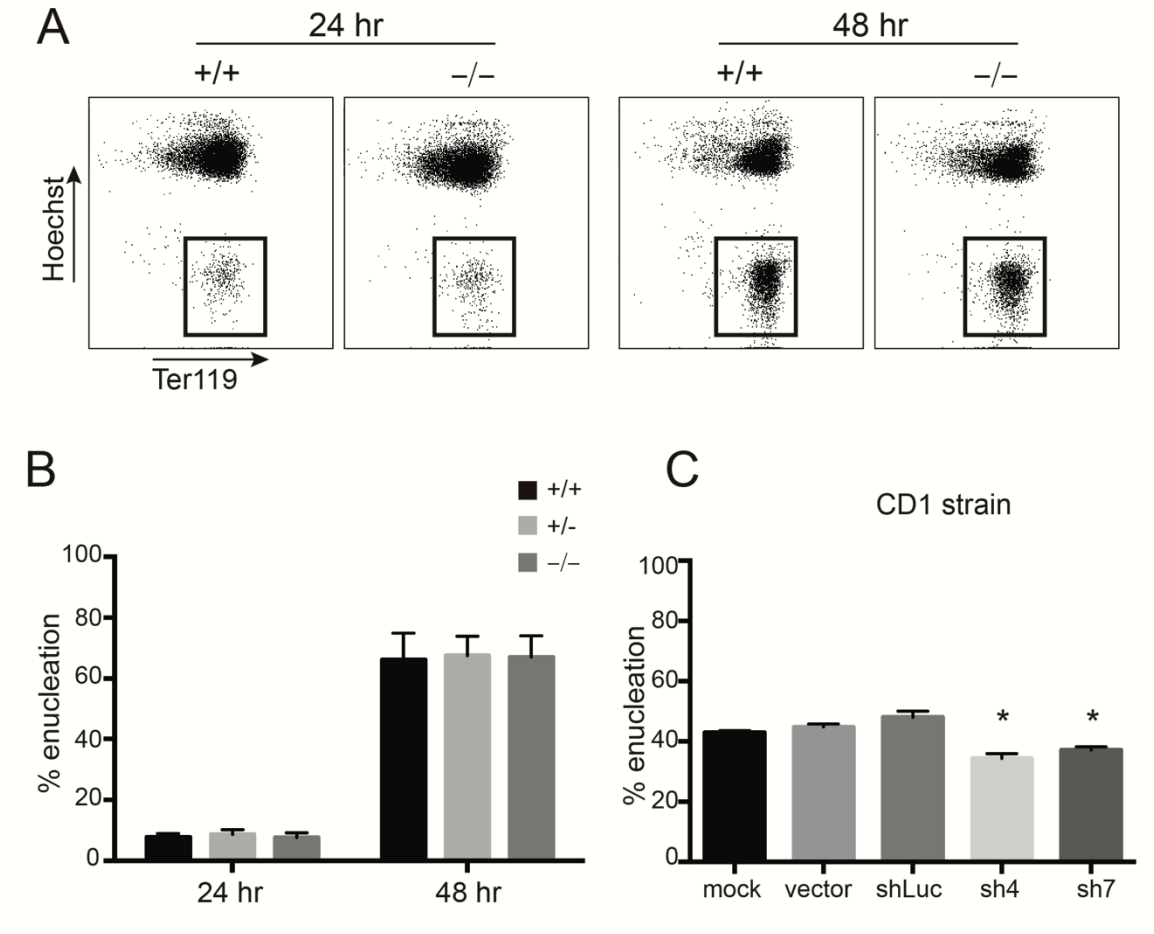


Figure 4.6. *Trim58*$^{-/-}$ and *Trim58*-deficient erythroblasts from mixed strain mice mature normally *in vitro.*

(A) Fetal liver erythroblasts isolated from E14.5 embryos (from heterozygous crosses) were cultured in erythroid expansion medium containing for 72 hours, then switched to maturation medium. Flow cytometry analysis is shown after 24 hours in maturation medium. Cells were stained with the erythroid maturation marker Ter119 and the cell permeable DNA dye Hoechst 33342 to assess percent (%) enucleated reticulocytes, indicated by the boxes. (B) Percent enucleated erythroblasts in fetal liver cultures at 24

and 48 hours of maturation. Data are expressed as mean ± SD for $Trim58^{+/+}$ (n=6), $Trim58^{+/-}$ (n=6), and $Trim58^{-/-}$ (n=10) embryos. (C) Enucleation of CD1 outbred strain erythroblasts. Data are presented as mean ± SD. Data are presented as mean ± SD. *$p<0.05$.

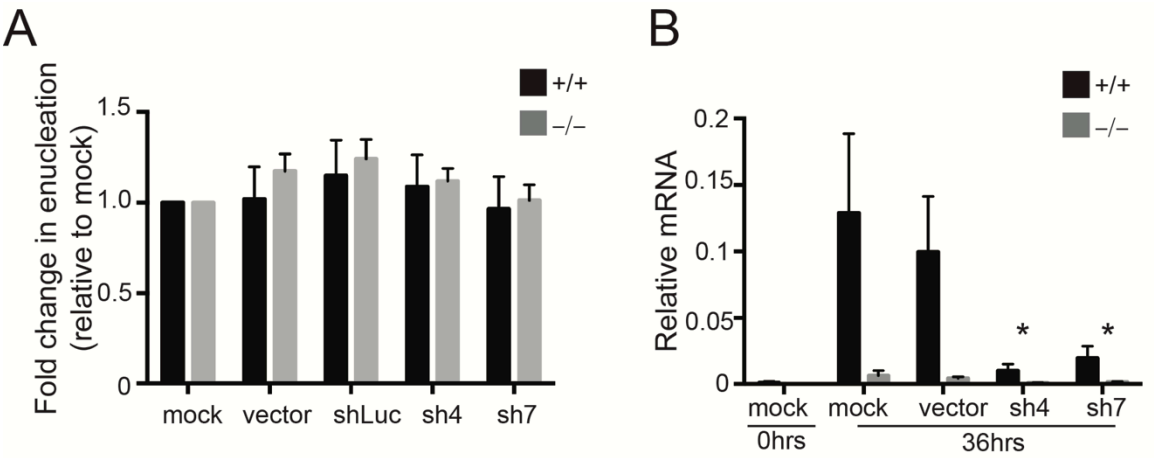


Figure 4.7. Trim58-directed shRNAs do not inhibit enucleation in fetal liver erythroblasts with mixed genetic background.

(A) Enucleation of mixed genetic background (Sv129, FVB, and C57Bl/6) *Trim58*$^{+/+}$ and *Trim58*$^{-/-}$ erythroblasts, with and without *Trim58*-directed shRNAs. Cells were analyzed after 36 hours of maturation. (B) *Trim58* mRNA levels determined by quantitative RT-PCR is shown in the left panel. Data are presented as mean ± SD. *$p<0.05$.

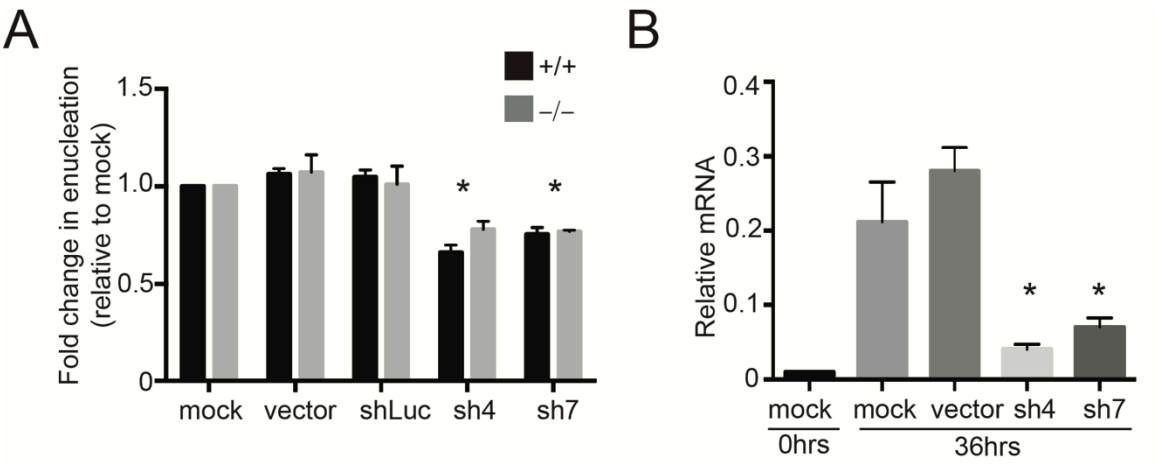


Figure 4.8. *Trim58* knockdown in wild type and *Trim58*$^{-/-}$ C57Bl/6 fetal liver erythroblasts reveals strain-specific off-target effect of shRNAs.

(A) Enucleation of C57Bl/6 *Trim58*$^{+/+}$ and *Trim58*$^{-/-}$ erythroblasts, with and without *Trim58*-directed shRNAs. (B) *Trim58* mRNA levels determined by quantitative RT-PCR is shown in the left panel. Data are presented as mean ± SD. *$p<0.05$. Cells were analyzed after 36 hours of maturation.

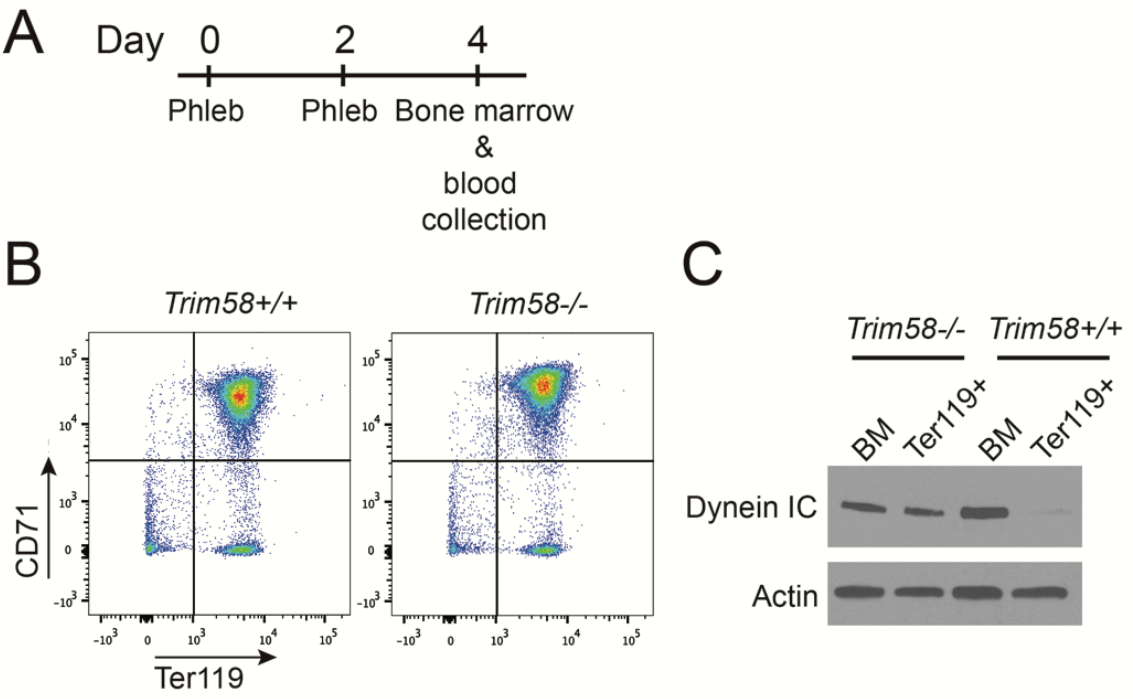


Figure 4.9. *Trim58*$^{-/-}$ erythroid cells in adult bone marrow accumulate dynein *in vivo*.

(A) Experimental scheme. Adult mice were subjected to phlebotomy (phleb), with 200μL blood removed on days 0 and 2. Peripheral blood and bone marrow were collected on day 4. (B) Ter119+ erythroblasts were isolated from bone marrow by immunomagnetic selection and analyzed by flow cytometry. (C) Western blot analysis for dynein intermediate chain (IC) in Ter119+ erythroblasts and whole bone marrow (BM) from *Trim58*$^{+/+}$ and *Trim58*$^{-/-}$ mice.

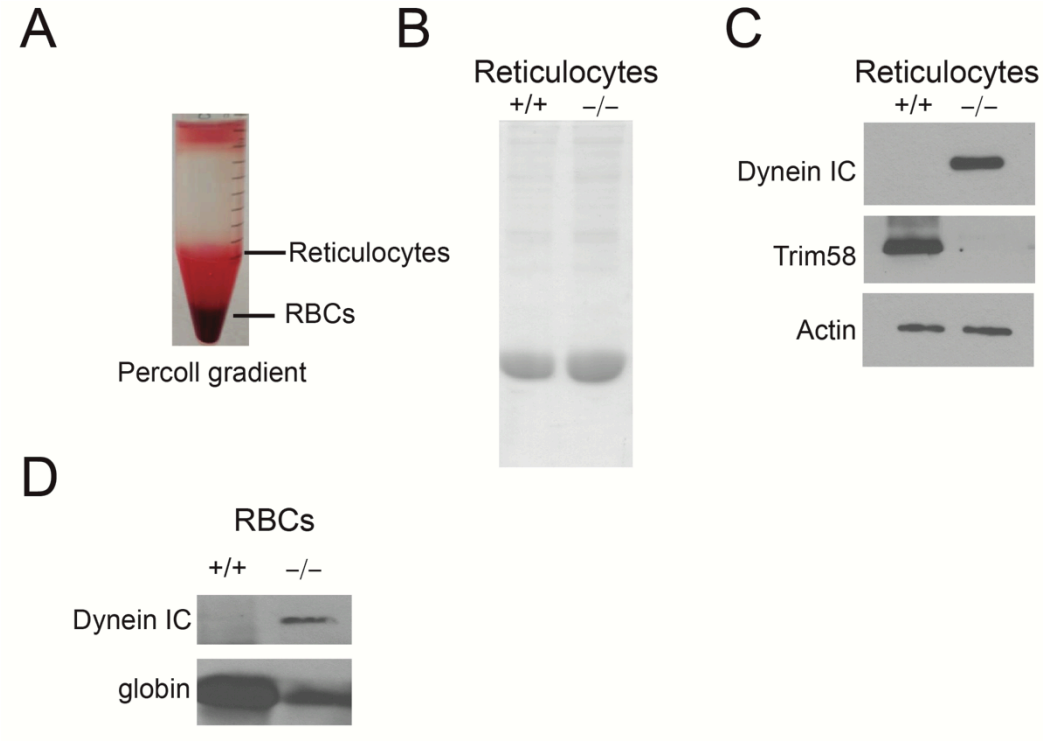


Figure 4.10. *Trim58*$^{-/-}$ erythroid cells in adult peripheral blood accumulate dynein *in vivo.*

(A) Whole blood fractionated by a Percoll gradient to isolate reticulocytes and red blood cells (RBCs). (B) Reticulocyte lysates analyzed by SDS-polyacrylamide gel electrophoresis and Coomassie blue silver stain. (C) Western blot analysis for dynein IC, Trim58 and β-actin in *Trim58*$^{+/+}$ and *Trim58*$^{-/-}$ reticulocytes. (D) Western blot analysis for dynein IC, Trim58 and β-actin in *Trim58*$^{+/+}$ and *Trim58*$^{-/-}$ RBCs.

Tables

	Sv129/FVB/Bl6		C57Bl/6	
	Trim58$^{+/+}$ (*n* = 3)	*Trim58*$^{-/-}$ (*n* = 3)	*Trim58*$^{+/+}$ (*n* = 12)	*Trim58*$^{-/-}$ (*n* = 6)
Hemoglobin	13.5±0.8	14.0±0.5	15.0±0.4	14.8±0.4
Hematocrit (%)	40.3±4.2	39.7±5.9	47.5±2.9	48.4±2.2
MCV	44.8±0.9	48.8±3.1	51.2±3.1	55.4±3.2 *
MCH	16.4±0.6	16.8±0.9	16.1±0.4	16.9±0.8
MCHC	33.9±5.5	35.7±6.1	31.2±1.7	30.5±1.3
RDW	16.9±0.7	17.3±0.5	17.6±1.0	18.3±0.7
RBC	10.0±0.2	9.7±0.3	9.3±0.4	8.7±0.4 *
Reticulocytes (%)	4.2±1.2	3.0±0.3	3.2±0.2	3.5±0.3
Platelets	771±167	549±94	830±103	794±56
WBC	1.9±1.0	2.6±1.2	7.0±2.0	5.2±2.5
Neutrophil (%)	16.4±4.5	17.2±1.6	13.7±3.6	12.2±3.4
Lymphocyte (%)	77.4±4.1	78.1±2.4	83.8±3.2	85.6±4.3
Monocyte (%)	4.9±2.4	4.2±1.2	2.4±1.2	2.1±0.6
Eosinophil (%)	1.0±1.3	0.4±0.4	0.1±0.1	0.1±0.1

Table 4.1. Complete blood counts of mixed background (Sv129/FVB/Bl6) and C57Bl/6 Trim58$^{+/+}$ and Trim58$^{-/-}$ mice (age 8 to 12 weeks).

The numbers of mice analyzed from each group are shown in parentheses. MCV, mean corpuscular volume (fL); MCH, mean corpuscular hemoglobin (pg); MCHC, mean corpuscular hemoglobin concentration (g/dL); RDW, red blood cell distribution width

(%); RBC, red blood cell number (x 10^6/μl); platelets (x 10^3/μL); WBC, white blood cells (x 10^3/μL). * $p < 0.05$, student's *t*-test.

Chapter 5 Genome editing recreates hereditary persistence of fetal hemoglobin (HPFH)

The first two chapters were inspired by studies of common human genetic variants associated with mild effects on RBC parameters. I became interested in studying rare mutations that caused more extreme phenotypes. Three additional factors influenced my investigation of fetal hemoglobin regulation, including 1) its clear relevance to human health; 2) the Weiss lab moved from Penn to St. Jude Children's Research Hospital, an institute with a long-standing priority of serving a large cohort of patients with sickle cell disease; and 3) the *Trim58*$^{-/-}$ mice were rederived at St. Jude and had reproducibly intact erythropoiesis. These studies were published in *Nature Medicine* (2016). Figures and Tables can be found at the end of the text.

Additional co-authors who contributed to this work:

Department of Hematology at St. Jude Children's Research Hospital: Yu Yao, Yong-Dong Wang, Kaitly J. Woodard, Chunliang Li, and Mitchell J. Weiss

RIKEN BioResource Center: Ryo Kurita and Yukio Nakamura

Weatherall Institute of Molecular Medicine, Oxford University: Jim R. Hughes

Pennsylvania State University: Ross C. Hardison

Division of Hematology at the Children's Hospital of Philadelphia: Gerd A. Blobel

Chapter summary

Disorders resulting from *HBB* (β-globin) gene mutations, mainly sickle cell disease (SCD) and β-thalassemia, become symptomatic postnatally as fetal γ-globin expression from two parologous genes *HBG1* and *HBG2* falls and adult β-globin increases, thereby shifting red blood cell (RBC) hemoglobin from the fetal (HbF, α2γ2) to adult form (HbA, α2β2). These disorders are alleviated when postnatal expression of fetal γ-globin is maintained. For example, in hereditary persistence of fetal hemoglobin (HPFH), a benign genetic condition, mutations attenuate γ-to-β switching, causing high-level HbF expression throughout life. Co-inheritance of HPFH with β-thalassemia or SCD mutations alleviates their clinical manifestations. Here we performed CRISPR-Cas9-mediated genome editing of human blood progenitors to mutate a 13-nucleotide *HBG1* and *HBG2* promoter sequence, thereby recapitulating a naturally occurring HPFH-associated mutation. Edited progenitors produced RBCs with increased HbF levels that were sufficient to inhibit pathological hypoxia-induced RBC morphologies of SCD. Our findings identify a potential DNA target for genome editing-mediated therapy of β-hemoglobinopathies.

Results

Typical HPFH mutations include heterozygous deletions or nucleotide (nt) substitutions within the extended β-globin locus (Forget, 1998). For example, the -175 T>C *HBG1* substitution de-represses γ-globin by creating a binding site for the

transcriptional activator TAL1 (Wienert et al., 2015). Other forms of HPFH may be associated with loss of cis elements that recruit γ-globin repressor proteins. We focused on a 13-nt deletion within the *HBG1* (Ag) promoter (-102 to -114, Figure 5.1) (Gilman et al., 1988). The deleted region contains a CCAAT box and direct repeat (DR), both of which recruit transcriptional repressor proteins (Liberati et al., 2001; Mantovani et al., 1989; Ronchi et al., 1995; Superti-Furga et al., 1988; Zhu et al., 2012). In heterozygous individuals approximately 30% of all RBCs express HbF protein (in individuals lacking the mutation <1% RBCs express HbF), which is potentially therapeutic for hemoglobinopathies, as SCD patients with this level of HbF are asymptomatic (Hoban et al., 2016). We reasoned that a DNA break induced by site-directed CRISPR/Cas9 mutagenesis, followed by error-prone non-homologous end joining (NHEJ) might recapitulate effects of the 13-nt-deleted HPFH mutation (Doudna and Charpentier, 2014). We used lentivirus to express mCherry, Cas9 and 2 different guide RNAs (gRNAs) targeting this region in the erythroblast cell line HUDEP-2, which expresses mainly HbA (Kurita et al., 2013) (Figure 5.1). HbF protein levels were undetectable in mock infected and Cas9 only expressing control cultures and rose to approximately 17% and 3% in cells expressing mCherry, Cas9 and gRNA-1 or gRNA-2, respectively (Figure 5.3A). Cells staining positive for HbF ("F-cells") increased from 2% to 46% with gRNA-1, and to 26% with gRNA-2 (Figure 5.2A). The %*HBG1/2* mRNA [$\gamma/(\gamma + \beta)$] was increased by both gRNAs; *HBB* mRNA expression was decreased by gRNA-1 (Figure 5.3B-C).

Induction of HbF by two adjacent non-overlapping gRNAs suggests that the effect is due to on-target mutations rather than to off-target mutations.

Next, we used the same mCherry/Cas9/ gRNA-encoding lentiviral vectors to edit peripheral blood $CD34^+$ hematopoietic stem/progenitor cells (HPSCs) from two healthy adults. Transduced HPSCs expressing mCherry were enriched by cell sorting and then induced to undergo erythroid differentiation. HbF protein rose from approximately 5% to 20% (Figure 5.2B), and F-cells increased from 18% to 54% in erythroid progeny of Cas9/gRNA-1-expressing $CD34^+$ cells (Figure 5.2C); gRNA-2 produced similar effects of lesser magnitude. Gene editing did not alter the expression of erythroid stage-specific maturation markers (Figure 5.4A-D). In FACS-purified a_4-$integrin^+$/$Band3^+$ late basophilic erythroblasts, gRNA-1 increased %*HBG1/2* mRNA from 4% to 35% and decreased absolute *HBB* mRNA by about 50% (Figure 5.4E-F). Increased γ-globin with decreased β-globin expression indicates reversal of the γ-to-β switch, which is controlled by competition of the corresponding genes for an upstream enhancer, termed locus control region (LCR) (Palstra et al., 2008).

To test potential therapeutic effects of this genome-editing approach, we edited patient-derived $CD34^+$ HSPCs from three SCD patients, induced erythroid differentiation, and examined the RBC progeny. Deoxygenated HbS forms rigid polymers that underlie the characteristic RBC morphology and pathophysiology of SCD, while increased levels of HbF inhibit this process. To test the effects of genome editing on HbS polymerization, we cultured SCD CD34+ cell-derived RBCs under hypoxia (2%

O_2). Approximately 30-40% of control and gene-edited cells matured to reticulocytes, late stage anucleate RBC precursors (Figure 5.5A). Consistent with previous reports, control (mock transduced or Cas9 alone) cultures contained relatively high percentages of F-cells (approximately 65%) (Akinsheye et al., 2011; Kidoguchi et al., 1978), although 25% of the cells within the entire population exhibited sickle morphology, presumably because HbF was absent in those cells or expressed at insufficient levels to prevent HbS polymerization (Figure 5.5B). In contrast, Cas9 + gRNA-1 expression increased F-cells to 90% and reduced sickle morphology to 4% (Figure 5.6A-B). Thus, targeting the -102 to -114 HPFH region inhibited HbS polymerization in cultured reticulocytes. Residual sickling in Cas9/gRNA-1-expressing RBCs could occur from lack of editing or mutations that do not induce HbF.

CRISPR/Cas9 creates targeted double-strand DNA breaks that are repaired via error-prone NHEJ, causing nt insertions and deletions (indels). g-globin is expressed from tandem homologous genes, *HBG2* and *HBG1* (Figure 5.7A). While the naturally occurring 13-nt HPFH deletion alters *HBG1*, *HBG2* contains an identical region that is mutated in different HPFH individuals (Forget, 1998; Stamatoyannopoulos, 2005). To characterize the mutations induced by Cas9/gRNA-1-expressing lentivirus in $CD34^+$ cells, we PCR-amplified a 431-nt region surrounding the predicted cleavage sites in *HBG*2 or *HBG1* and deep-sequenced the products. In 3 biological replicate experiments, the indel incidence in all transduced cells was 56, 65 and 77%, with equal mutation rates in *HBG1* and *HBG2* (not shown). Approximately half of the mutations were identical to

the -102 to -114 HPFH deletion (not shown). Most likely, the 13-nt deletion predominates because the Cas9 cleavage site is flanked by 8-nt tandem repeats that facilitate microhomology-mediated end-joining (Truong et al., 2013). To identify mutations associated with g-globin induction, we purified low-, intermediate-, and high-HbF erythroblasts and analyzed the promoter regions by deep sequencing (Figure 5.7B). In 4 independent experiments, the 13-nt HPFH deletion was enriched in HbF-high cells ($p<0.05$) (Figure 5.8a and Figure 5.7C).

Transduction of $CD34^+$ cells with Cas9/gRNA-1 lentivirus produced numerous indels smaller than 13 nt (Figure 5.8a and Figure 5.7C), although their low frequencies impeded efforts to determine the effects on HbF expression. Thus, we performed clonal studies by expressing gRNA-1/Cas9 in $CD34^+$ cells transiently via DNA electroporation followed by methylcellulose culture to analyze single cell-derived burst-forming unit-erythroid (BFU-E) colonies (Figure 5.9 and Figure 5.10). Of 344 colonies screened from 3 experiments, 30 had on-target mutations, with 16 containing at least one 13-nt HPFH deletion (Figure 5.10). All BFU-E colonies analyzed were mosaic for *HBG1/2* mutations, reflecting editing over several progenitor divisions. *HBG1* and *HBG2* were mutated equally. The mutation frequencies correlated moderately with %*HBG1/2* mRNA levels in each colony (Figure 5.9) (r^2=0.41, $P < 0.0001$). Several small indels within the CCAAT box and/or DR were associated with elevated γ-globin expression in one or more colonies (Figure 5.10). We also derived HUDEP-2 cell clones after transient Cas9/gRNA-1 expression, including three lines with small (1-4 nt) bi-allelic *HBG1/HBG2* CCAAT box

mutations and strong γ-globin induction (Figure 5.11A-C and Figure 5.12). HUDEP clone 6 (Figure 5.11C) and BFU-E colony 4 (Figure 5.10) contained single-nt *HBG1* and *HBG2* insertions that preserved the core CCAAT box element and altered the DR, but was not associated with elevated γ-globin expression.

Simultaneous double-stranded DNA cleavage at gRNA-1 recognition sites in the *HBG2* and *HBG1* promoters could result in NHEJ-mediated joining of the two ends with loss of the intervening 5.2 kb. Indeed, we identified this deletion in several genome edited HUDEP-2 clones (Figures 5.13 and 5.14). We determined the frequency of this event in transiently edited cells using quantitative PCR (qPCR) and fluorescence in situ hybridization (FISH) (Figure 5.13A). The 5.2-kb deletion was not detected in HUDEP-2 cells after electroporation of 2 mg Cas9/gRNA-1 DNA plasmid, but occurred in about 20% of *HBG1/2* alleles after 10 mg DNA transfection (Figure 5.13B-C). The deletion was not detected by qPCR in human $CD34^+$ HSPCs electroporated with 10 mg Cas9/gRNA-1 plasmid (Figure 5.14). Lastly, deep sequencing of $CD34^+$ cells after lentiviral delivery of Cas9/gRNA-1 showed no indels at the top 15 bioinformatically predicted off-target sites (Table 5.1).

The -114 γ-globin promoter CCAAT box and overlapping DR element likely mediate postnatal γ-to-β globin switching by recruiting transcriptional repressors in a developmentally regulated fashion (Forget, 1998; Stamatoyannopoulos, 2005). Candidate CCAAT box binding proteins include COUP-TFII (NR2F2) (Liberati et al., 2001), NF-Y

(CP-1, CBF) (Liberati et al., 2001; Zhu et al., 2012), NF-E3 (Mantovani et al., 1989; Ronchi et al., 1995), CDP (Mantovani et al., 1989; Superti-Furga et al., 1988), and C/EBP (Superti-Furga et al., 1988); the DR element, binds nuclear hormone receptors TR2 and TR4 (Tanabe et al., 2002). Although the molecular triggers of γ-to-β globin switching are not fully defined, the current study shows that disruption of the -114 *HBG1/HBG2* CCAAT box/DR region via gene editing partially reverses the switch in adult-type erythroid cells and delineates further the cis elements involved. Importantly, γ-globin induction can occur via small (< 13 nt) NHEJ-associated mutations caused by transient editing and independent of microhomology-mediated end-joining, which is preferentially utilized during S-phase and may not occur efficiently in hematopoietic stem cells (Truong et al., 2013).

Chapter commentary

Genome editing technologies to manipulate hematopoietic stem cells have fueled innovative strategies for treating β thalassemia and SCD, including correction of the SCD mutation by homology-directed DNA repair (Hoban et al., 2015), reactivation of γ-globin via forced promoter-LCR looping (Deng et al., 2014) or disruption of the repressor gene *BCL11A* via NHEJ (Canver et al., 2015). These approaches are untested in patients. Here we present an additional approach; CRISPR-Cas9-mediated disruption of an *HBG1/2* region associated with a benign human condition (HPFH) induces HbF expression to

potentially therapeutic levels, similar to those achieved via forced DNA looping or Cas9/gRNA-mediated disruption of a *BCL11A* erythroid enhancer. Future studies are now required to optimize editing of the γ-globin CCAAT box/DR site in human hematopoietic stem cells and minimize potentially harmful off-target mutations. Overall, our study provides proof-of-principle for a potential approach to treat common hemoglobinopathies by genome editing.

Figures

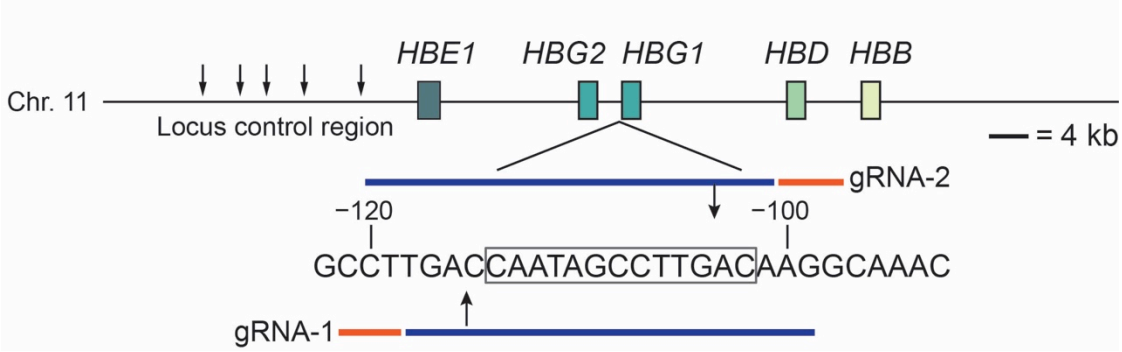


Figure 5.1. Extended β-globin locus.

Extended β-globin locus showing β-like genes as boxes. Small arrows mark DNase I hypersensitive sites within the locus control region, an upstream enhancer. A region of the *HBG1* promoter is shown numbered according to position upstream of the transcription start with the 13-nt HPFH deletion boxed. Guide RNA spacer sequences are blue and PAM motifs (NGG) are orange; gRNA-1 and gRNA-2 are complementary to the sense and antisense strands, respectively. Large arrows show predicted Cas9 cleavage sites.

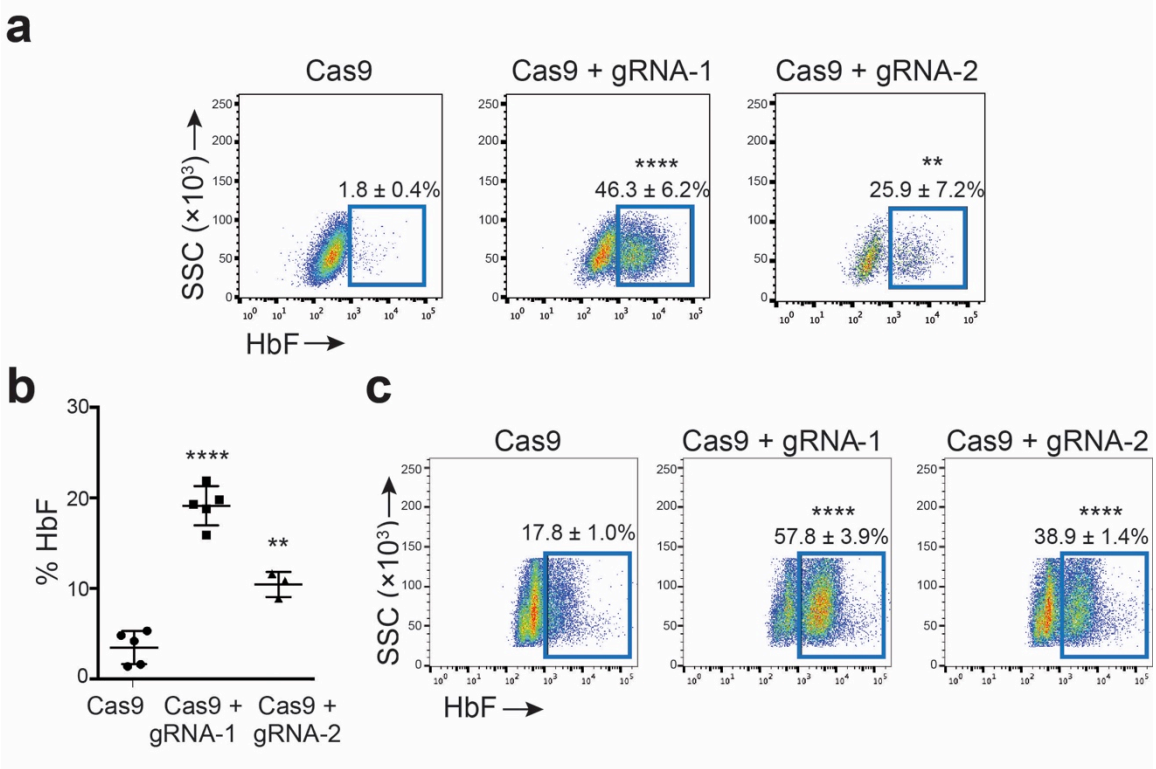


Figure 5.2. Genome editing of the *HBG1* and *HBG2* promoters increases erythroid fetal hemoglobin (HbF) levels.

(**a**) Representative flow cytometry plots showing HbF^+ immunostaining HUDEP-2 cells 5d after transduction with Cas9 ± gRNA-1 or gRNA-2 lentivirus. Numbers indicate mean ± standard error (SE) from four independent experiments.

(**b**) Normal human $CD34^+$ cells transduced with lentivirus encoding Cas9 ± gRNA-1 or gRNA-2 were cultured for 21d in erythroid cytokines, then analyzed for hemoglobin (Hb) protein by HPLC. %HbF = [HbF/(HbA + HbF) × 100]. Each dot represents a separate

experiment performed with $CD34^+$ cells from the same donor. On-target editing rates of *HBG1*/*HBG2* in three experiments were 56%, 65% and 77%.

(**c**) HbF^+ erythroblasts, derived as described for panel (b). Numbers indicate mean ± SE from three experiments.

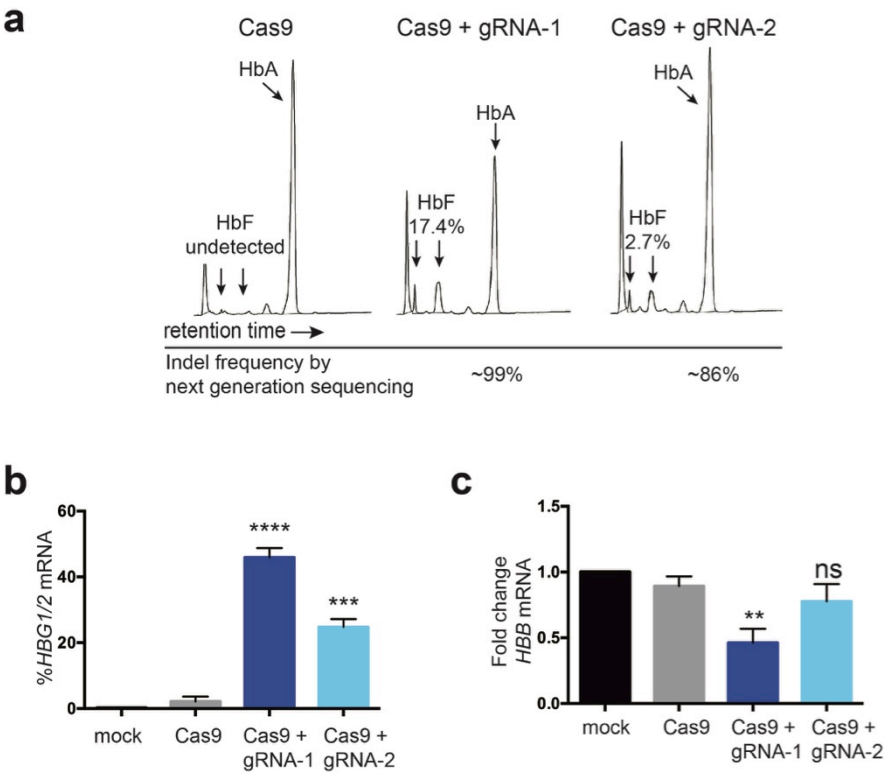


Figure 5.3. Genome editing of the HBG1 and HBG2 promoters reverses the γ-to-β globin switch in HUDEP-2 cells.

HUDEP-2 cells were transduced with lentivirus expressing mCherry and Cas9 ± gRNAs. mCherry+ cells were purified by FACS, cultured for 3 days, then analyzed. (a) Quantification of HbA (α2β2) and HbF (α2γ2) proteins in undifferentiated HUDEP-2 cell cultures expressing Cas9 ± gRNA-1 or -2 by high performance liquid chromatography (HPLC). Shown below are indel allele frequencies, expressed as percentages of all alleles, as determined by PCR amplification and next generation sequencing using primers depicted in Supplementary Fig. 4a. Coverage for each nt

position sequenced was over 97,000x. (b) Quantitative real-time PCR analysis showing %*HBG1/2* (γ-globin) mRNA [γ/(γ + β)] in gene-edited and control cells. n = 3, shown as mean value ± standard error (SE). (c) *HBB* (β-globin) mRNA levels in genome-edited and control cells, expressed as fold change vs. mock-transduced cells. n = 3, mean value ± SE. ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ by unpaired t-test.

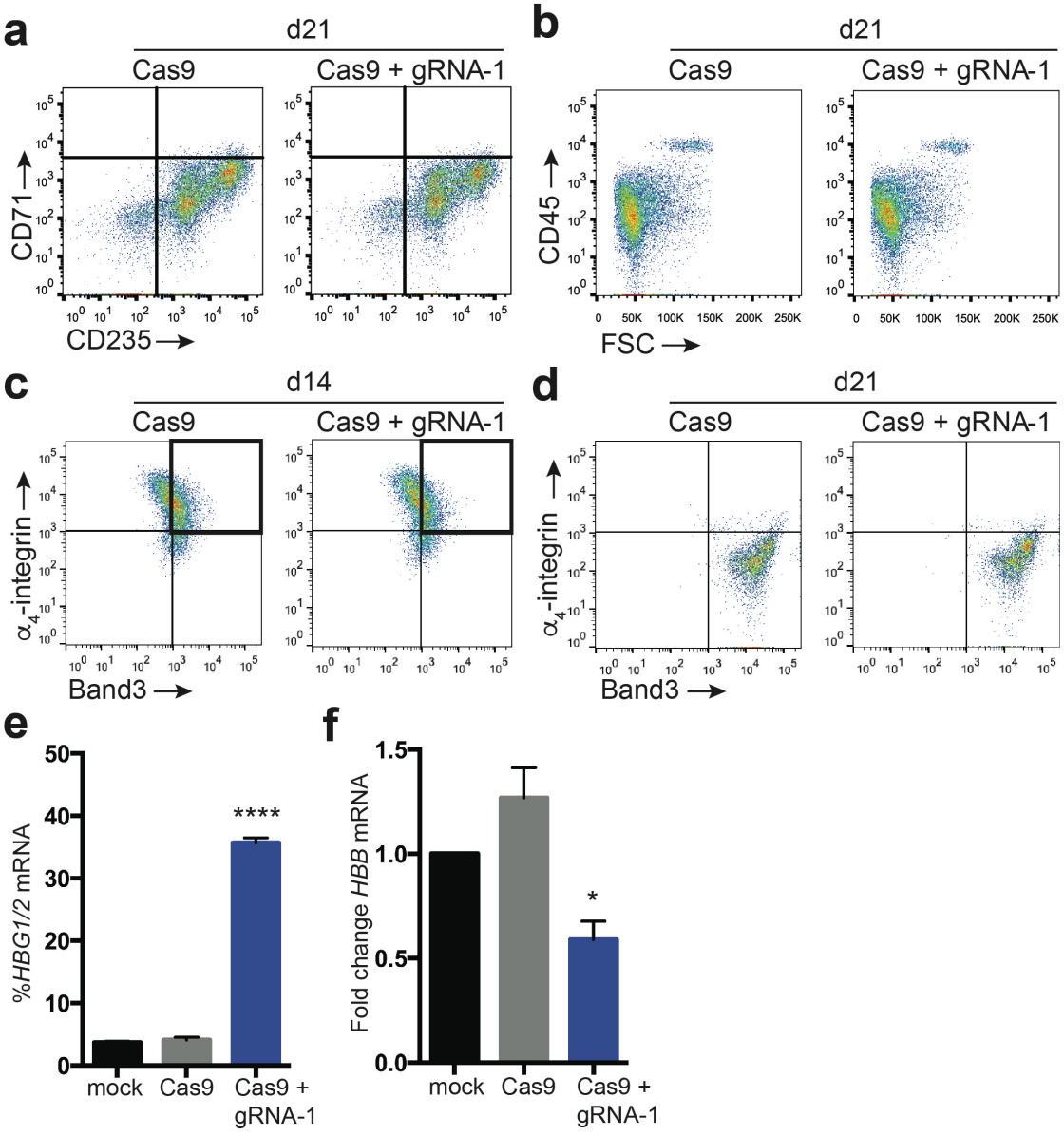


Figure 5.4. Genome editing of the HBG1 and HBG2 promoters in CD34+ cells reverses the γ-to-β globin switch in erythroid progeny without altering differentiation or maturation.

CD34+ hematopoietic stem and progenitor cells (HSPC) from a healthy donor were transduced with lentivirus expressing Cas9 or Cas9/gRNA-1, mCherry+ cells were FACS

enriched, cultured under conditions favoring erythroid differentiation and analyzed by flow cytometry for developmental stage specific markers. The plots shown are representative of three experiments. (a) Erythroid maturation markers CD71 and CD235 on culture day 21. (b) CD45 expression on culture day 21. (c-d) α4-integrin downregulation and Band3 upregulation between d14 and d21 of erythroid maturation. Cells in gates outlined in black were isolated by FACS and analyzed by qRT-PCR for %*HBG1/2* (γ-globin) mRNA in panel (e) and (f) *HBB* (β-globin) mRNA levels, relative to mock-transduced cells. Overall indel frequencies for three independent experiments in $CD34^+$ cells were 56, 65, and 77%. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ by unpaired student's t-tests.

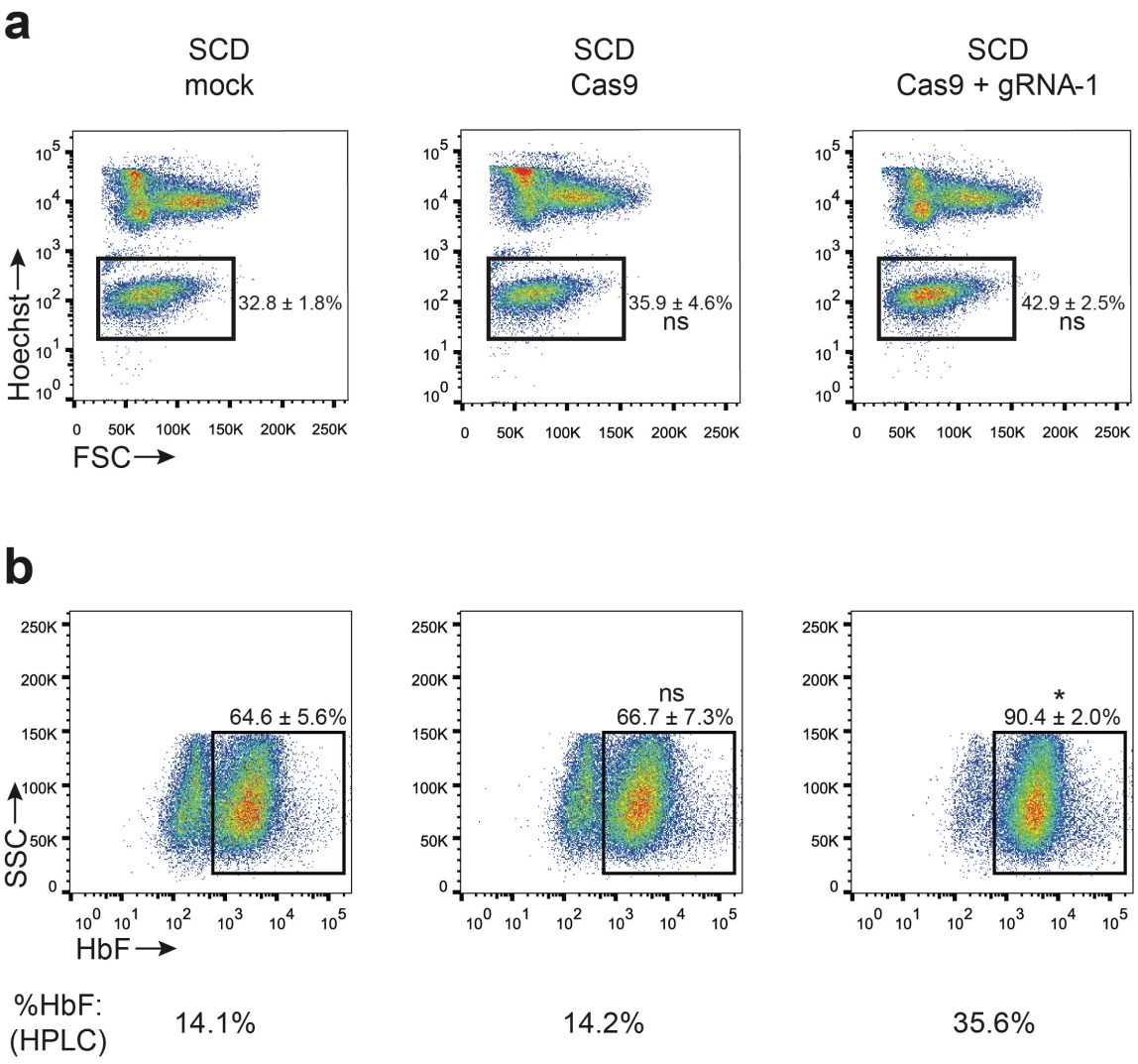


Figure 5.5. Effects of Cas9 + gRNA-1 on SCD erythroblasts.

Peripheral blood CD34+ cells from three SCD patients with HbSS genotype were transduced lentivirus expressing Cas9 or Cas9/gRNA-1, enriched by FACS for mCherry fluorescence, and cultured under conditions that favor erythroid maturation. After 21 days of culture, cells were analyzed by flow cytometry. Representative plots from three experiments are shown with mean values ± SE. (a) Hoechst staining to distinguish

nucleated erythroblasts from anucleate reticulocytes (black boxes). (b) HbF immunostaining cells ("F-cells", black boxes). Numbers in the flow cytometry plots for panels (a) and (b) indicate mean ± SE for three independent experiments. Numbering below the plots in panel (b) indicates HPLC determination of %HbF for one experiment. * $P < 0.05$, ns = not significant, unpaired student's t-tests.

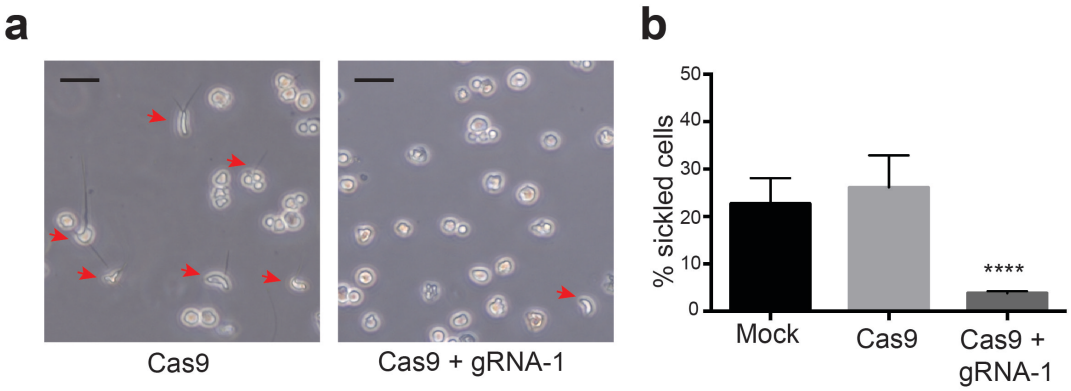
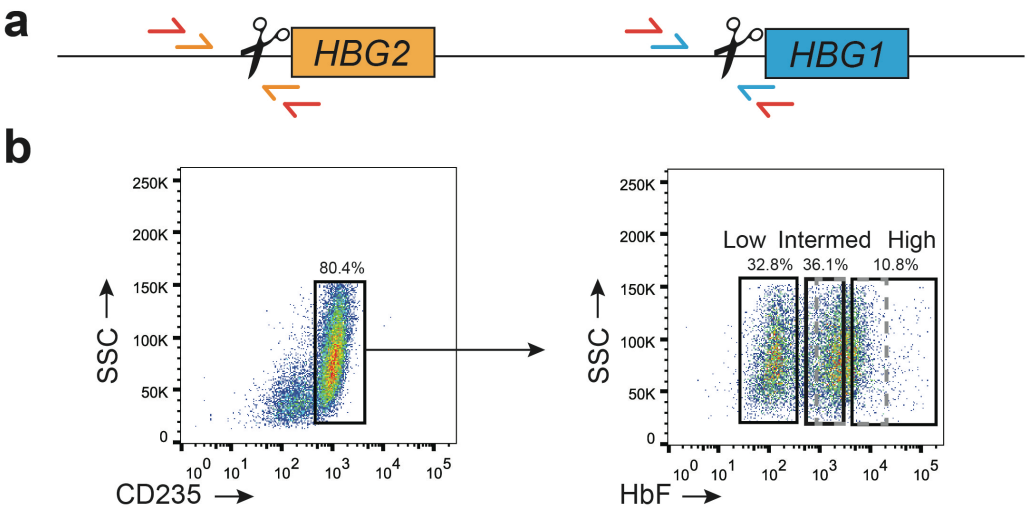


Figure 5.5. Genome editing inhibits sickling *in vitro*.

(**a**) $CD34^+$ cells from an SCD (HbSS) patient were transduced with lentivirus expressing Cas9 ± gRNA-1, differentiated into RBCs, and cultured in 2% O_2.Red arrows denote cells with sickle-like morphology. Original magnification, 200×. Size bars indicate 20 μm.

(**b**) Quantification of hypoxia-induced sickled cells depicted in panel (a). Mean ± SE from three experiments using $CD34^+$ cells from three different SCD donors (> 1,000 cells scored per experiment).



c

	Allele frequency (%)											
	HbF: Low				Intermediate				High			
	rep1	rep2	rep3	rep4	rep1	rep2	rep3	rep4	rep1	rep2	rep3	
GCCTTGCCTTGACCAATAGCCTTGACA												
GCCTTGCCTTGAC----------A	10.7	12.5	31.1	16.6	35.3	40.5	37.2	45.1	54.8	61.2	46.6	
GCCTTGCCTTGA-CAATAGCCTTGACA	4.3	2.7	10.8	9.6	7.7	4.9	8.6	7.8	5.9	4.2	9.3	
GCCTTGCCTTGA--AATAGCCTTGACA	1.9	1.8	6.9	3.6	3.7	3.7	6.8	4.8	2.5	2.9	6.2	
GCCTTGCCTTGA---ATAGCCTTGACA	1.7	1.5	6.2	7.4	3.2	3.4	5.2	6.6	1.8	2.5	2.9	
GCCTTGCCTTGA----TAGCCTTGACA	0.7	0.6	1.9	0.4	2.0	1.8	2.8	2.4	3.9	4.2	4.0	
GCCTTGCCTTGAcCCAATAGCCTTGACA	1.7	0.8	1.8	2.0	1.1	1.1	1.0	1.3	1.0	0.5	2.2	
GCCTTGCCTTG-CCAATAGCCTTGACA	0.5	0.2	1.6	0.6	1.4	0.7	1.8	2.4	1.3	1.0	1.8	
G----------CCAATAGCCTTGACA	0.3	0.2	1.3	0.8	0.8	0.6	2.0	1.9	0.7	0.8	1.6	
GCCTTG------CCAATAGCCTTGACA	0.5	0.3	1.5	0.8	1.0	0.4	1.1	1.8	1.5	0.7	0.9	

Figure 5.6. HBG1 and HBG2 promoter mutation analysis of CD34+ HSPC-derived erythroblasts transduced with Cas9/gRNA-1 lentivirus.

(a) PCR amplification design for the target region, not drawn to scale. The scissors represent the gRNA-1 cleavage site upstream of both *HBG1* and *HBG2* genes. The half arrows denote PCR primers used to amplify the target regions. The red primers amplify both *HBG1* and *HBG2* nonspecifically. The orange and blue primers specifically amplify

HBG1 or *HBG2*, respectively. (b) Cell sorting scheme to isolate low, intermediate, and high HbF-expressing CD235+ erythroblasts that were analyzed further according to main Figure 5.8 and in panel (c) of this figure. Rep4 was gated into two populations, HbF-low and HbF-high, as indicated by the dashed gray box. The indicated percentages refer to rep1 shown in Figure 5.8. Cells were fractionated at d16 of erythroid culture. (c) Results of replicates (rep 1-4) for the experiment in this Figure and in main Fig. 2a. The DNA surrounding the predicted Cas9-gRNA cleavage sites in *HBG1/HBG2* were amplified by PCR (red primers in panel (a)) and analyzed by next generation sequencing. Overall indel frequencies for unfractionated cells were 65% (rep2) and 77% (rep3). The wild type sequence for HBG1/HBG2 is shown at the top left, above the 9 most common mutant alleles. The gray box shows the 13-nt HPFH deletion and tandem 8-nt repeats are shown by blue half arrows. The CCAAT box is shown in red. The 9 most common mutated alleles are shown below, with dashes indicating deleted nucleotides and lower-case letters indicating insertions. Mutant allele frequencies are expressed as percentages of all alleles, shown at right.

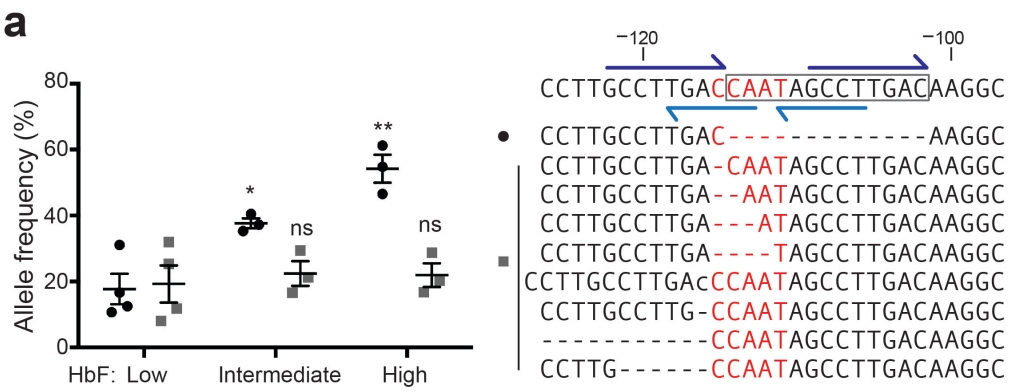


Figure 5.7. Spectrum of γ-globin-inducing mutations caused by Cas9 and gRNA-1.

(**a**) Normal adult $CD34^+$ cells were edited with Cas9/gRNA-1 lentivirus, differentiated into RBCs, FACS-purified according to HbF immunostaining intensity (low, intermediate or high) and analyzed for on-target mutations. The wild type sequence is shown on the top left with the 13-nt HPFH deletion boxed and the CCAAT box in red. Dark blue half arrows show flanking 8-nt repeats; light blue show the DR element. The top nine mutant alleles (of more than 40 total indels identified) are shown below. Dashes indicate nucleotide deletions and lower-case letters insertions. In the graph at right, black dots denote the 13-nt HPFH deletion, which occurred at the highest frequency; gray squares show the combined frequencies of the eight next common mutations. Each symbol represents an independent experiment. * $P < 0.05$, ** $P < 0.01$ by unpaired *t*-test.

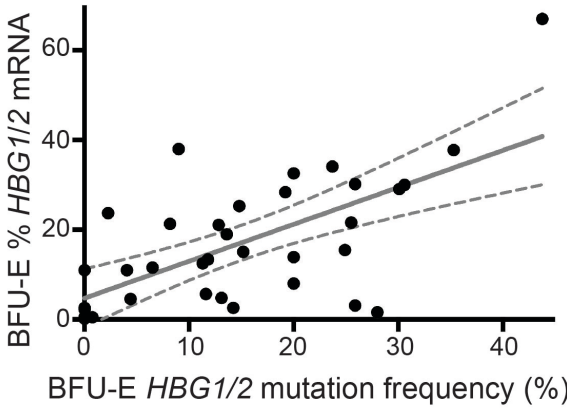
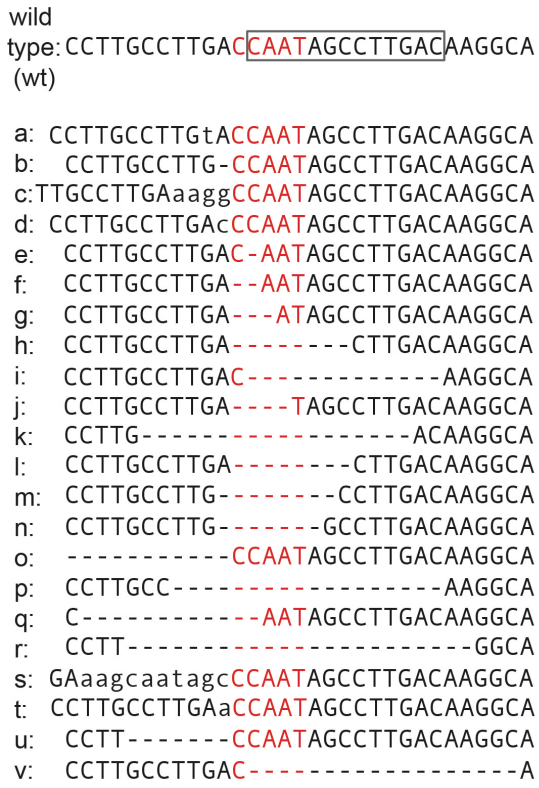


Figure 5.8. Mutation rates in BFU-E colonies correlates positively with γ-globin expression.

$CD34^+$ cells were electroporated with Cas9/gRNA-1/GFP expression plasmids. GFP^+ cells were FACS purified and seeded into methylcellulose. Burst forming unit-erythroid (BFU-E) colonies were analyzed for globin mRNAs and *HBG1*/*HBG2* mutations. All colonies were mosaic for mutations; the total *HBG* [(*HBG1* + *HBG2*)/2 × 100] mutation frequency for each colony is plotted against %*HBG1/2* mRNA [γ / (γ + β)]. Regression analysis shows best-fit line as solid gray ($y = 0.82x + 4.8$, $r^2 = 0.41$, $P < 0.0001$, $n = 35$ colonies from three experiments) and 95% confidence intervals as dashed gray.

Colony #	*HBG1* mutation frequency (%)	*HBG2* mutation frequency (%)	Total mutation frequency (%)	%*HBG1/2* mRNA
1	wt	a: 47.3	23.6	34.1
2	b: 38.1 c: 6.7	b: 7.2	26.0	30.2
3	b: 11.2	b: 50.0	30.6	30.0
4	d: 56.1	wt	28.0	1.6
5	wt	e: 13.0	6.5	11.6
6	wt	f: 18.0	9.0	38.0
7	g: 22.2 f: 5.6 h: 1.8	wt	14.8	25.3
8	wt	i: 16.4	8.2	21.3
9	i: 47.0 d: 10.6	d: 2.6	30.1	29.1
10	i: 6.7	f: 19.4	13.0	4.8
11	wt	j: 49.3 i: 1.7	25.5	21.6
12	i: 21.9 g: 1.8	wt	11.8	13.4
13	i: 8.7	wt	4.4	4.6
14	c: 3.3	c: 11.9 i: 7.9	11.6	5.7
15	wt	k: 22.7	11.4	12.5
16	f: 1.5	wt	0.8	0.5
17	i: 16.7	e: 13.0 d: 8.6	19.2	28.4
18	wt	f: 51.8	25.9	3.1
19	d: 11.3	g: 12.5 k: 3.5	13.6	19.0
20	b: 28.0 i: 2.7 l: 2.2	d: 7.1	20.0	8.0



Colony #	*HBG1* mutation frequency (%)	*HBG2* mutation frequency (%)	Total mutation frequency (%)	%*HBG1/2* mRNA
21	b: 28.5	wt	14.2	2.6
22	wt	i: 33.0 d: 4.0 m: 1.8 n: 1.3	20.0	13.9
23	o: 9.4 p: 8.5 q: 2.2	p: 9.1 q: 1.2	15.2	15.1
24	j: 63.1	b: 4.1 i: 2.4 j: 1.0	35.3	37.8
25	g: 33.2 r: 1.8	f: 14.8	24.9	15.5
26	i: 4.5	wt	2.2	23.7
27	i: 18.8 s: 14.3 b: 4.1 n: 1.6	t: 33.1 b: 7.6 i: 6.7 u: 1.4	43.8	67.0
28	wt	i: 23.8 n: 1.9	12.8	21.1
29	b: 1.4	i: 6.8	4.1	11.0
30	wt	v: 29.7 i: 10.4	20.0	32.6
31-35	wt	wt	0.0	0.2-11.0

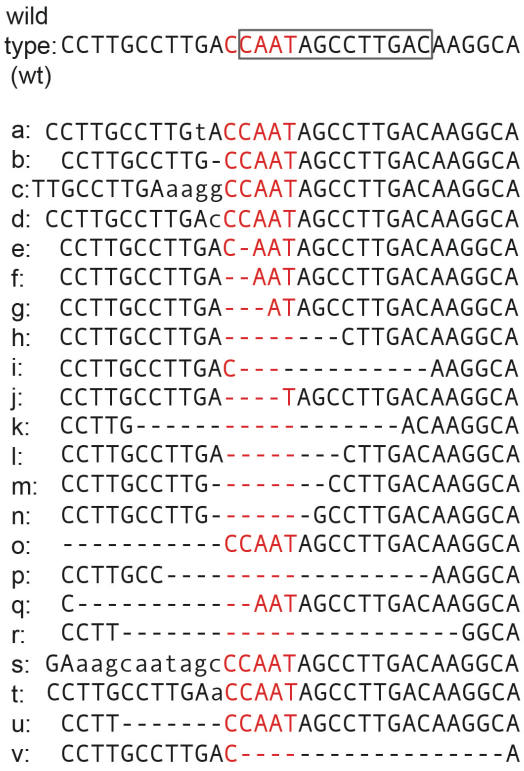


Figure 5.9. Analysis of burst forming unit-erythroid (BFU-E) colonies derived from transiently edited CD34+ HSPCs.

CD34+ HSPCs were electroporated with 10 µg gRNA-1/Cas9/GFP plasmid. GFP+ cells were enriched by FACS and cultured in methylcellulose with erythroid-promoting cytokines. Of 344 individual colonies evaluated for *HBG1* and *HBG2* mutations, we detected mutations in 30. The wild type sequence is depicted at the top right, with the 13-nt HPFH deletion boxed. Specific mutations identified in each colony are shown below the wild type sequence. Dashes denote deleted nucleotides, and lower-case denotes

inserted bases. Frequencies for all detected mutations are expressed as percentages of all alleles.

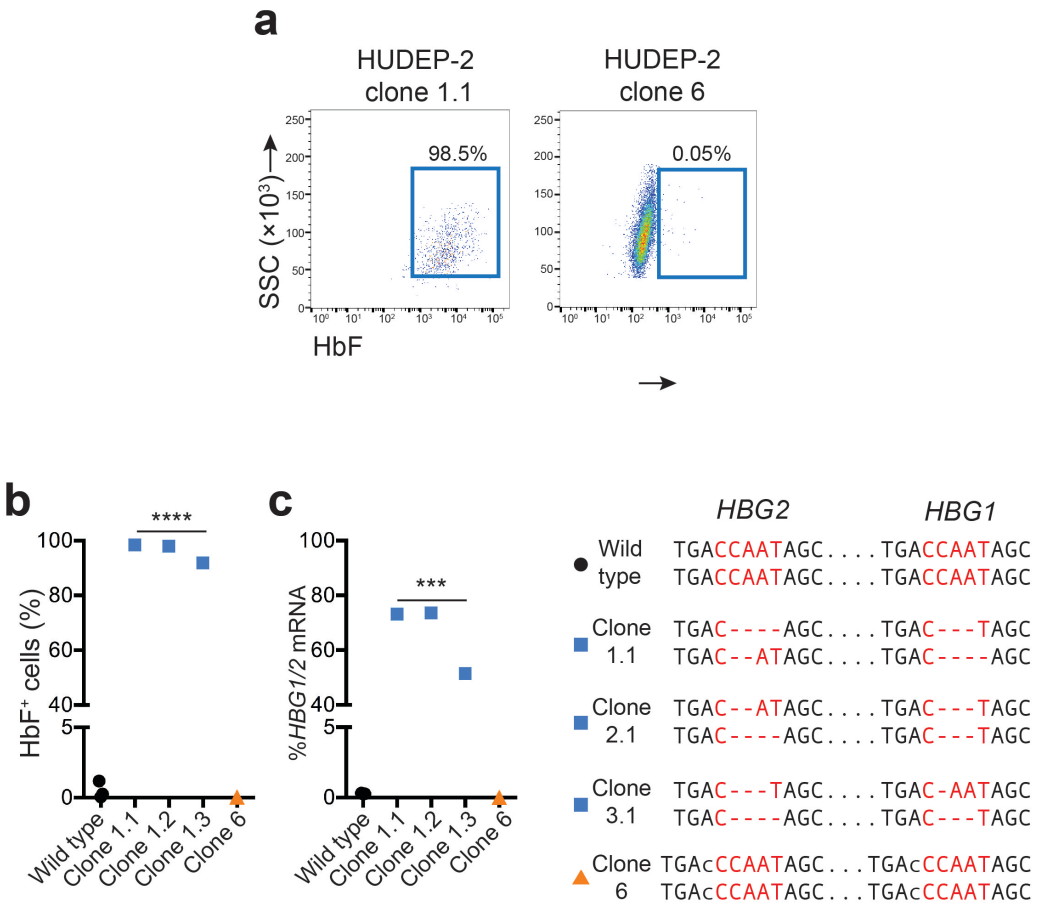
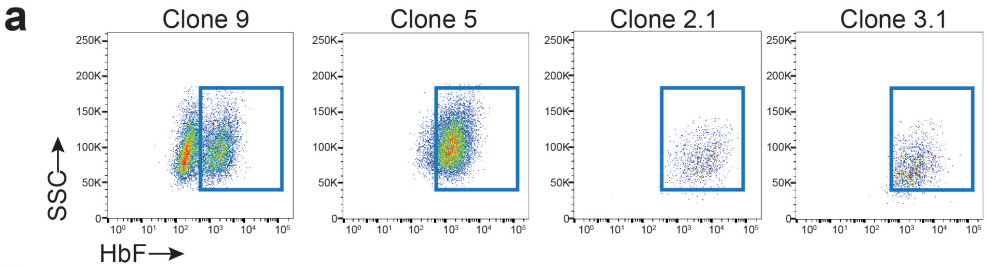


Figure 5.10. HUDEP-2 clones with CCAAT box mutations demonstrate high HbF expression.

(**a**) HUDEP-2 cells were electroporated with Cas9/gRNA-1/GFP plasmid and cloned. HbF immunostaining is shown for two representative clones with different mutations (see also panel (d) and Figure 5.12).

(**b-c**) Characterization of genome edited HUDEP-2 clones. The *HBG1* and *HBG2* genotypes corresponding to each clone is shown on the right, according to the convention used in panel (a); all clones are homogenous for the indicated mutations. *** $P < 0.001$, **** $P < 0.0001$, ns = not significant by unpaired *t*-test.



b

Clone	*HBG2*	*HBG1*	%HbF⁺ cells	%*HBG1/2* mRNA
Wild type	TTGACCAATAGCCTTGACAA . . . . . . TTGACCAATAGCCTTGACAA	TTGACCAATAGCCTTGACAA TTGACCAATAGCCTTGACAA	0.25	0.27
1.1	TTGAC----AGCCTTGACAA . . . . . . TTGAC--ATAGCCTTGACAA	TTGAC---TAGCCTTGACAA TTGAC----AGCCTTGACAA	98.5	73.2
2.1	TTGAC--ATAGCCTTGACAA . . . . . . TTGAC----AGCCTTGACAA	TTGAC---TAGCCTTGACAA TTGAC---TAGCCTTGACAA	98.0	73.6
3.1	TTGAC---TAGCCTTGACAA . . . . . . TTGAC----AGCCTTGACAA	TTGAC-AATAGCCTTGACAA TTGAC---TAGCCTTGACAA	91.9	51.5
4.1	TTGAC---TAGCCTTGACAA . . . . . . TTGA----------	TTGAC-AATAGCCTTGACAA -----CAATAGCCTTGACAA	98.4	not done
5	TTGAC----------AA TTGAC----------AA		96.2	89.7
6	TTGAcCCAATAGCCTTGACAA . . . . . . TTGAcCCAATAGCCTTGACAA	TTGAcCCAATAGCCTTGACAA TTGAcCCAATAGCCTTGACAA	0.05	0.02
7	TTGAC----------AA(53.7%) TTGAC-AATAGCCTTGACAA(18.9%)	TTGAC---TAGCCTTGACAA(16.6%)	98.5	not done
8	TTGAC----------AA(18.7%)	TTGAC----------AA(68.3%) TT--------GCCTTGACAA (4.5%) TT-------AGCCTTGACAA (4.5%) TT----ATAGCCTTGACAA (2.7%)	89.8	not done

Figure 5.11. Analysis of genome-edited HUDEP-2 clones.

HUDEP-2 cells were electroporated with DNA plasmid encoding Cas9/gRNA-1/GFP. Single GFP+ cells were sorted into 96-well plates after 24h and expanded. Clones were analyzed for *HBG1/2* mutations and HbF immunostaining. Of 47 clones examined, 70% were mosaic for *HBG1* and *2* mutations. Four mosaic clones were sub-cloned to generate clones 1.1, 2.1, 3.1 (also shown in main Figure 5.11) and 4.1 (shown here). (a) Clone 9 is

a mosaic with a bi-modal distribution of HbF immunostaining cells. (b) Analysis of clonal populations for *HBG1* and *HBG2* mutations, HbF immunostaining and %*HBG1/2* (γ-globin) mRNA [γ/(γ + β)]. Dashes indicate deleted nt. Lower-case letters indicate base insertions. Clones 5-6 contain the indicated mutations in all cells and were not sub-cloned. Clones 7-8 are shown as examples of mosaic clones. Clones 4.1 and 5 are heterozygous and homozygous for a 5.2-kb deletion between Cas9/gRNA-1 cleavage sites in the *HBG1/2* promoters. Clone 6 had a normal copy number in the same qPCR assay but may have a deletion hundreds of nt long that prevents PCR/sequencing detection, yet preserves the qPCR primer/probe sites.

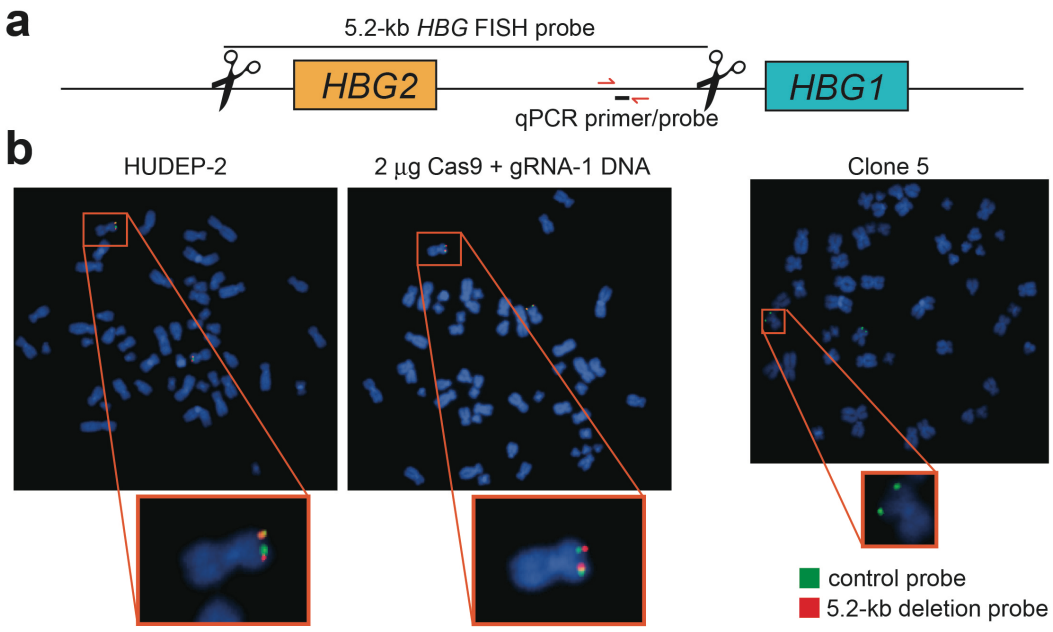


Figure 5.12. Deletion analysis of Cas9/gRNA-1 edited HUDEP-2 cells.

Cells were electroporated with DNA plasmid encoding Cas9/gRNA-1/GFP. After 24h, GFP+ cells were FACS-purified and analyzed. (a) Map of *HBG2* and *HBG1* showing cleavage sites for Cas9/gRNA-1 (scissors). Fluorescence in situ hybridization (FISH) and quantitative real-time PCR (qRT-PCR) were used to detect deletion of the intervening 5.2-kb region. (b) FISH analysis of edited HUDEP-2 cells using the 5.2-kb probe depicted in panel (a). The control probe detects the downstream HBB gene. No 5.2-kb deletions were detected in 25 metaphase cells analyzed from each group. Loss of 5.2-kb FISH probe signal was observed for a clone homozygous for the 5.2-kb deletion (clone 5, Supplementary Fig. 6).

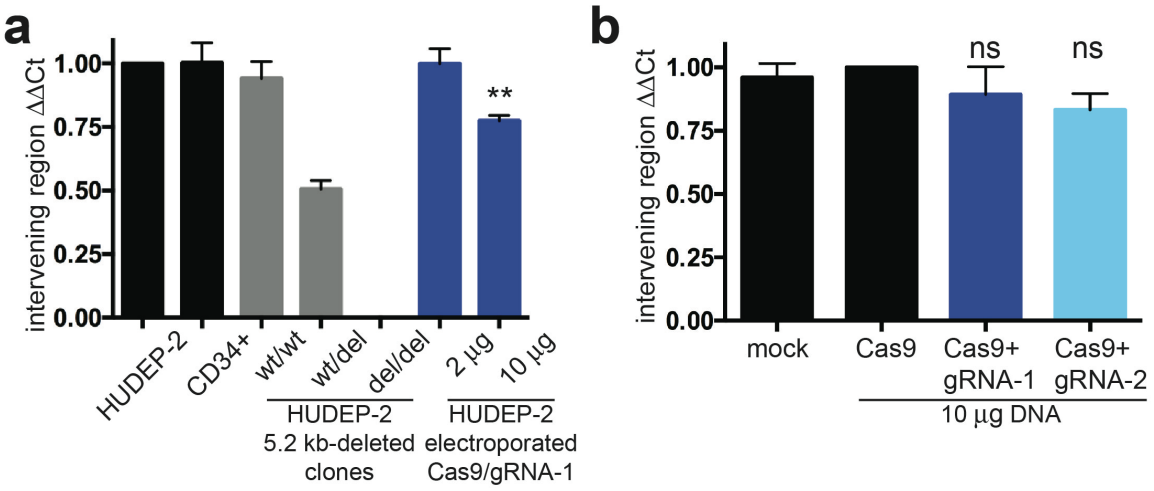


Figure 5.13. Deletional analysis of Cas9/gRNA-1 edited HUDEP-2 and CD34+ cells.

(a) qRT-PCR analysis for the 5.2-kb deletion in edited HUDEP-2 cells. Black bars indicate wild type control cells. Gray bars indicate HUDEP-2 lines that are heterozygous (wt/del) or homozygous (del/del) for the deletion, corresponding to clones 4.1 and 5, respectively, described in Supplementary Fig. 6. Blue bars indicate HUDEP-2 cells electroporated with 2 µg or 10 µg Cas9/gRNA-1 plasmid. Representative results are shown from two experiments; values are normalized to wild type HUDEP-2 cells. (b) Deletion analysis after transient Cas9/gRNA-1 expression in human CD34+ cells. Dark and light blue bars represent HSPCs electroporated with 10 µg DNA encoding Cas9/gRNA-1/GFP or Cas9/gRNA-2/GFP, respectively.

Tables

Target	Target Sequence	Identity	Mis-matches	Strand	Chromosome Position (hg38)	Type	Nearest Gene		Coverage	
							5' (kb)	3' (kb)	mock	Cas9/ gRNA-1
	CTTGTCAAGGCTATTGGTCA AGG									
1	CTTGGCTTGGCTATTGGTCA TGG	17/20	3	-	15:77307168-77307190	Intronic	PEAK1		480K	220K
2	CTTGTCAGGGCTATTGCTTA AGG	17/20	3	-	20:6584634-6584656	Intergenic	CASC20 (56.2)	BMP2 (183.0)	400K	300K
3	CTAGTCAAGGCTGTTTGTCA GGG	17/20	3	+	20:19275552-19275574	Intronic	SLC24A3		92K	69K
4	CTTGTCAGGGCTGTTGGTCG AGG	17/20	3	-	5:137486079-137486101	Intronic	SPOCK1		300K	98K
5	CTTGTCCAGGCTGCTGGGCA CGG	16/20	4	-	1:47708454-47708476	Intergenic	FOXD2 (267.8)	TRABD28 (51.9)	195K	185K
6	CTTGTCCAGGCTCCTGGGCA CGG	16/20	4	-	1:47708482-47708504	Intergenic	FOXD2 (267.8)	TRABD28 (51.9)	195K	185K
7	CTTGTCCAGGCTCCTGGGCA CGG	16/20	4	-	1:47708510-47708532	Intergenic	FOXD2 (267.8)	TRABD28 (51.9)	195K	185K
8	CTTGTCCAGGCTCCTGGGCA CGG	16/20	4	-	1:47708538-47708560	Intergenic	FOXD2 (267.8)	TRABD28 (51.9)	195K	185K
9	CTTGTCCAGGCTCCTGGGCA CGG	16/20	4	-	1:47708566-47708588	Intergenic	FOXD2 (267.8)	TRABD28 (51.9)	195K	185K
10	CTTGTCCAGGCTCCTGGGCA CGG	16/20	4	-	1:47708594-47708616	Intergenic	FOXD2 (267.8)	TRABD28 (51.9)	195K	185K
11	CTTGTCAAGGCTTTATGTAA GGG	16/20	4	+	1:173221950-173221972	Intergenic	TNFSF4 (14.6)	LOC100506023 (13.1)	360K	290K
12	CTAGTCAAGGCAATTAGGCA AGG	16/20	4	+	1:240226990-240227012	Intronic	FMN2		22K	33K
13	TTTGTCTAGGATATTGCTCA GGG	16/20	4	+	10:22210628-22210650	Intergenic	EBLN1 (0.65)	LOC100130992 (41.4)	400K	460K
14	CTTGTGATGGCGATGGGTCA TGG	16/20	4	+	10:25181498-25181520	Intronic	GPR158		380K	380K
15	CTGGGAAAGGCTATTGGTAA TGG	16/20	4	+	10:96412345-96412367	Intronic	TLL2		230K	94K

Table 5.1. Predicted gRNA-1 off-target sites assessed by deep sequencing.

Potential off-target sites were predicted using the Sanger off-target prediction tool: http://www.sanger.ac.uk/htgt/wge/find_off_targets_by_seq. The top 15 off-target sites predicted by nucleotide homology are shown, with mismatched nucleotides in red. To evaluate potential off-targeting at these loci, human $CD34^+$ cells were transduced with lentivirus encoding Cas9, gRNA-1, and mCherry. FACS was performed to enrich for mCherry-positive, infected cells, which were then cultured in erythroid-promoting cytokines. Genomic DNA from erythroblasts from untransduced cells (mock) and cells infected with Cas9/gRNA-1 was extracted and used as template to amplify the DNA

surrounding the top 15 candidate sites shown. Coverage for each amplicon is shown at right, and no mutant alleles were detected for any amplicon.

For a full list of predicted off-targets:

http://www.sanger.ac.uk/htgt/wge/crispr/1077773870

Chapter 6 Conclusions and Future Directions

6.1 The role of Trim58 in terminal erythropoiesis

In the first set of studies, I investigate the role of an erythroid-specific E3 ubiquitin ligase, Trim58, which is highly induced during terminal erythropoiesis. We found that Trim58 mediates proteasomal degradation of dynein in maturing erythroblasts, and our analysis of Trim58-deficient cells suggested that Trim58 was necessary for erythroid enucleation. We extended our initial studies by disrupting the *Trim58* gene in mice to investigate its role in dynein degradation and erythropoiesis *in vivo*. The gene targeting strategy removed exon 3, which encodes a coiled-coil domain predicted to participate in protein oligomerization (Koliopoulos et al., 2016; Sanchez et al., 2014; Streich et al., 2013). Ablation of exon 3 maintains the *Trim58* mRNA reading frame and therefore could produce an internally deleted protein lacking only the coiled-coil motif (Trim58 ΔCC, Figure 2S). However, homozygous mutant erythroblasts contained only trace amounts of *Trim58* ΔCC mRNA and no detectable Trim58 protein. Thus, the gene targeting likely produced a null allele in RBCs. Although it was possible to express Trim58 ΔCC in heterologous cells, the mutant protein was unable to degrade dynein, demonstrating that the coiled-coil domain is necessary for ubiquitin ligase activity, as observed for other Trim protein family members (Koliopoulos et al., 2016; Sanchez et al., 2014; Streich et al., 2013). Erythroid cells from *Trim58*$^{-/-}$ mice expressed dynein at abnormally high levels, which resulted in mildly altered MCV and RBC number, but no anemia, hemolysis or altered recovery after phenylhydrazine-induced hemolytic anemia.

Therefore, loss of *Trim58* and failure to eliminate dynein are largely dispensable for erythropoiesis.

6.1.1 Reconciling *in vitro* and *in vivo* systems

Previously, we showed that *Trim58* shRNAs delayed the kinetics of enucleation in cultured primary erythroblasts. However, our *in vivo* results demonstrate this process to be normal in *Trim58*$^{-/-}$ erythroblasts. Moreover, *Trim58* shRNAs delayed enucleation similarly in C57Bl/6 and CD1 (but not mixed Sv129/FVB/C57Bl/6) Trim58$^{+/+}$ and *Trim58*$^{-/-}$ erythroblasts, indicating that this shRNA effect was strain-specific and independent of *Trim58* silencing. shRNAs can cause off-target effects by interacting with mRNAs via partial complementarity. However, this mechanism is unlikely since multiple non-overlapping *Trim58* shRNAs inhibited enucleation (Jackson et al., 2006).

More plausible explanations include shRNA-induced interferon response (Jackson and Linsley, 2004) or saturation of endogenous RNA processing pathways by high levels of virally expressed shRNAs with consequent inhibition of cellular microRNAs. In hepatocytes, this effect occurs partly via inhibition of exportin-5, which mediates nuclear export of shRNAs and microRNAs (Grimm et al., 2006). Late stage erythroblasts may be particularly susceptible to oversaturation of microRNA/shRNA processing proteins as the proteome complexity progressively declines. In fact, numerous aspects of erythroid maturation, including enucleation, are regulated by endogenous microRNAs (Rouzbeh et al., 2015; Zhang et al., 2011; Zhao et al., 2010), even as exportin-5 level declines

(Hattangadi et al., 2014). Moreover, most shRNAs expressed to sufficient levels, including controls, inhibit enucleation of cultured erythroblasts (Paralkar et al., 2014).

6.1.2 Are erythroid cells sensitive to generalized miRNA inhibition?

Our unexpected results have led us to hypothesize that specific components of the miRNA machinery may be dose-limiting in terminally differentiating erythroblasts. A miRNA-sequencing study in erythroid cells at different stages of development demonstrated that most miRNAs are downregulated during erythropoiesis. While we predict that most miRNAs with this pattern of expression would remain unaffected by limited reserves of miRNA machinery, several important miRNAs are highly induced during erythroid maturation. miR-144 and miR-451 are transcribed from a bicistronic locus, and both are highly expressed in mature erythroid cells (Yu et al., 2010). In fact, miR-451 is the most abundant miRNA expressed, accounting for 58% of all miRNAs at the Ter119^{+} stage (Zhang et al., 2011). We plan to assess the effects of varying levels of shRNA expression on both endogenous production and downstream processing of miR-144 and miR-451. Interestingly, miR-451 processing is Dicer-independent (Cheloufi et al., 2010; Cifuentes et al., 2010), so altered levels of mature miR-451 would help pinpoint the rate-limiting factor pathway. While further studies are required to define the exact mechanism by which *Trim58* shRNAs inhibit erythroblast enucleation, our observations indicate that the effects of any shRNAs on erythroid maturation must be interpreted with caution and validated through alternate lines of investigation. This

particular line of investigation becomes clinically important with the proposed gene therapy approach using shRNAs targeting *BCL11A* specifically in erythroid cells to induce fetal hemoglobin (see Chapter 1.5) (Guda et al., 2015).

6.1.3 The role of dynein degradation in hematopoiesis

Our results in Chapters 3 and 4 are consistent in demonstrating that Trim58 mediates degradation of dynein, a multi-subunit molecular motor complex that transports organelles and proteins along microtubules, usually in a minus direction (toward the centromere). Dynein participates in essential cellular processes including mitosis and organelle positioning and is necessary for the proliferation of erythroid precursors and their subsequent enucleation (Kobayashi et al., 2016). Thus, it is puzzling as to why Trim58 expression is induced during late stage erythropoiesis when dynein is necessary for these processes. Perhaps the unique maturation of this cell type is sensitive to dynein levels, which Trim58 helps to fine tune. In this regard, dynein levels may regulate erythroid MCV by influencing specialized cell divisions of late stage erythroid precursors, which are associated with shortened G1 phase and progressively reduced cell volume (Dolznig et al., 1995; Grebien et al., 2005). This process is modulated by the levels of Cyclin D3 (Sankaran et al., 2012) and Cyclin A2 (Ludwig et al., 2015) whose corresponding genes (*CCND3* and *CCND2A*), like *TRIM58*, were discovered by GWAS to regulate erythroid MCV (Ganesh et al., 2009; Kamatani et al., 2010; van der Harst et al., 2012). During thrombopoiesis, *Trim58* expression may influence the rates of dynein-

dependent proplatelet formation and maturation to regulate platelet size and number. In any case, the timing and levels of Trim58 induction during erythroid and megakaryocyte maturation must be regulated precisely so as to preserve dynein for general and specialized cellular processes. In agreement, retroviral expression of Trim58 in erythroid precursors removed dynein prematurely and induced cell death (Thom et al., 2014).

While premature depletion of dynein is likely deleterious to erythropoiesis, the current study shows that abnormally elevated dynein does not impair erythroid maturation, although in principle its excessive activity could alter the trafficking of cellular organelles. However, we found no evidence for erythroblast enucleation defects or abnormal distribution of lysosomes and mitochondria in *Trim58*$^{-/-}$ reticulocytes (not shown). During erythroblast enucleation, the microtubule network partially collapses and segregates with the nascent reticulocyte (Thom et al., 2014). Shortly thereafter, the UPS degrades residual microtubules (Chasis et al., 1989), and we were unable to detect them by α-tubulin immunostaining of circulating *Trim58*$^{+/+}$ or *Trim58*$^{-/-}$ mouse reticulocytes (not shown). Thus, while dynein is expressed at abnormally high level in *Trim58*$^{-/-}$ reticulocytes and RBCs, it is probably inactive without microtubules and not toxic.

From the work in Chapter 3, we formulated two models to explain the regulation of dynein protein levels during enucleation (Figure 3.17): 1) dynein suppression was necessary to create a molecular motor imbalance, allowing the plus end-directed actions of kinesins to polarize the nucleus away from the MTOC; and 2) Trim58-mediated dynein degradation destabilized microtubules and promoted their detachment from both

the cell cortex and nuclear membrane (i.e. microtubule cage collapse) (Hendricks et al., 2012). The former model was supported by studies performed in other cell types, including muscles and neurons (McKenney et al., 2010; Splinter et al., 2010; Tanenbaum et al., 2011; Tsai et al., 2010; Wilson and Holzbaur, 2012). In addition, dynein coordinates mitosis by mediating chromosome movements, spindle organization, and spindle positioning (Sharp et al., 2000; Howell et al., 2001;Varma et al., 2008). Data in Chapter 4 suggest these mechanisms are not dysregulated sufficiently disrupt erythropoiesis. Dynein excess, as observed in *Trim58*$^{-/-}$ erythroid cells, could potentially delay these processes and mitotic progression, consequently resulting in mild changes in RBC size and number.

Does Trim58 regulate megakaryocyte and platelet function?

Of note, Trim58 is also expressed in megakaryocytes (Pimkin et al., 2014). The platelet count trended downward in *Trim58*$^{-/-}$ mice, although this did not reach statistical significance. Dynein-dependent sliding of antiparallel microtubules is required for the extension of proplatelets from megakaryocytes and subsequent platelet formation (Bender et al., 2015; Patel et al., 2005). Dynein is also involved in platelet activation by inhibiting contraction of the marginal band, a ring-shaped network of microtubules on the periphery of platelets (Diagouraga et al., 2014). Therefore, *Trim58* could potentially influence platelet production and size by modulating dynein expression in these cells. In agreement with the latter, GWAS links *TRIM58* to circulating platelet counts. Our results

suggest that rather than platelet count, *Trim58* may affect the function of megakaryocytes and platelets. Alpha and dense granules are membrane-bound organelles, containing platelet-activating factors including von Willebrand factor and serotonin, which are actively translocated along microtubules and sorted into proplatelets (Richardson et al., 2005). These granules are secreted upon platelet activation to promote stability of a clot at a site of blood vessel injury. Trim58 may, therefore, function in megakaryocytes to limit granule delivery. Further studies are required to investigate whether *Trim58*$^{-/-}$ mice exhibited altered dynein levels in megakaryocytes and platelets and the consequent effects on formation and function of the latter.

6.1.4. Lessons from GWAS

Genome wide association studies identify nucleotide polymorphisms that associate with traits of interest and typically ascribe causality to altered function or expression of the nearest gene. However, causal genetic variants that impact cell traits frequently modify transcriptional regulatory elements (enhancers), which can potentially influence the expression of distant genes (Ulirsch et al., 2016). Therefore, GWAS findings must be verified by functional studies to identify and characterize the relevant involved gene. Approximately 75 candidate genes that regulate RBC traits have been identified by GWAS (Ganesh et al., 2009; Kamatani et al., 2010; van der Harst et al., 2012). Follow-up studies have elucidated new aspects of erythroid physiology, including regulation of cell size (*CCND3* and *CCND2A*) (Ludwig et al., 2015; Sankaran et al.,

2012), erythroblast proliferation (*SH2B3*) (Giani et al., 2015), RNA splicing (*RBM38*) (Ulirsch et al., 2016), and fetal hemoglobin production (*BCL11A* and *MYB*) (Bauer et al., 2013; Lettre et al., 2008; Sankaran et al., 2008a). The current study demonstrates that loss of *Trim58* expression alters MCV and RBC number, consistent with previous GWAS linking the human *TRIM58* gene to the same RBC traits, and providing a potential mechanism by showing that Trim58 is a ubiquitin ligase that facilitates dynein degradation. Of note, the effect magnitude of *Trim58* ablation in mice was approximately 27- and 67-fold greater for MCV and RBC number, respectively, compared to the effect magnitudes associated with human *TRIM58* variant alleles (see Chapter 2), which likely cause subtle alterations in protein expression or function.

Why complete loss of *Trim58* in mice and a subtle alteration of *TRIM58* in humans produce similar mild effects on RBCs (and perhaps platelets) is unknown. One potential explanation may be that non-redundant requirements for *TRIM58* are greater in human erythropoiesis compared to mice. Gene expression in mice and human erythroblasts are surprisingly discordant, with potential functional consequences (An et al., 2014; Pishesha et al., 2014). For example, mutations in the protein secretory transporter gene *SEC23B* cause congenital dyserythropoietic anemia in humans, while *Sec23b* knockout mice exhibit normal erythropoiesis (Khoriaty et al., 2014). It is also possible that enhanced *TRIM58* activity conferred by human genetic variation, with consequent lowering of dynein levels, exerts more profound effects on erythroid maturation than loss of *Trim58* with abnormal dynein retention. In this case, complete

loss of *Trim58* in mice would produce relatively small effects on RBC number and MCV (as observed), while effects of similar magnitude in the opposite direction could result from more subtle alterations in *TRIM58* caused by naturally occurring allelic variants. Regardless, our studies confirm human GWAS by demonstrating that Trim58 regulates RBC traits. However, Trim58 is not essential for generating these cells in mice, but rather, serves to fine-tune their terminal maturation, possibly by restricting dynein levels.

It is also possible that Trim58 function is redundant with other E3 ubiquitin ligases or other protein degradation pathways. Germline knockout of *Trim58* might allow gene expression changes during development permitting $Trim58^{-/-}$ mice to compensate for gene dosage. This effect has been reported when comparing phenotypes observed with RNAi-mediated knockdown or morpholino versus genetic ablation (Rossi et al., 2015). To begin to address these questions, we performed microarray transcriptome profiling, comparing $Ter119^{+}$ fetal liver erythroblasts from $Trim58^{+/+}$ and $Trim58^{-/-}$ E14.5 embryos (not shown). Outside of Trim58, we did not identify any genes with significant, differential expression between the two groups. These data support the former hypothesis, where the UPS is highly critical, and overlapping pathways exist at steady state to protect against loss of a single element. Another Trim protein, Trim10, previously shown to mirror the expression pattern of Trim58 in erythroid cell lines (Blaybel et al., 2008; Harada et al., 1999). Like $Trim58^{-/-}$ mice, $Trim10^{-/-}$ mice also do not show obviously aberrant erythropoiesis (personal communication, J. Downing). Future work

will investigate this by interbreeding mouse models, as it is possible that multiple components of the UPS pathway need to be mutated to affect the erythroid compartment.

6.2. Genome editing the γ-globin promoters

In Chapter 5, we tested the hypothesis that CRISPR-Cas9 could generate mutations in the *HBG1* and *HBG2* promoters that would recapitulate the effects of a 13-nt HPFH deletion in human hematopoetic stem and progenitor cells (HSPCs). We chose to target this locus mainly for the technical advantage of the mutation being small. We predicted that targeting the mutation would require a single gRNA and would also be generated at a higher frequency. The observed deletion efficiency with a pair of gRNAs is inversely related to the size of the large deletion (Canver et al., 2014). We found that lentiviral-mediated editing of $CD34^+$ HSPCs and an erythroid cell line, HUDEP-2, indeed caused phenotypes of HPFH. Importantly, edited erythroblasts underwent normal erythroid differentiation and maturation, suggesting that defects in or delayed erythropoiesis was not responsible for γ-globin induction. Furthermore, we explored the feasibility of targeting the γ-globin promoters in proof-of-concept experiments using SCD patient-derived HSPCs. Our data show that genome editing reduced *in vitro* sickling by about 75% and support -102 to -114 of the *HBG* promoters as potential DNA targets for therapeutic genome editing for β thalassemia and SCD.

6.2.1 Therapeutic genome editing for SCD

In the context of therapy development, genome-editing technologies to manipulate hematopoietic stem cells have dramatically fueled innovative strategies for treating β thalassemia and SCD. Homology-directed recombination (HDR)-mediated correction of the SCD mutation is perhaps the most theoretically direct approach, involving an extra-chromosomal template to serve as the wild type allele. The main limitation of this approach is the relatively inefficiency in HPSCs. Hoban et al. recently reported *in vitro* mutation correction rates of 20% in $CD34^+$ HSPCs using a targeting zinc-finger nuclease and integration-deficient lentivirus encoding template for repair (Hoban et al., 2015). However, after transplanting the human cells into a xenograft mouse model, the corrected allele frequency plummeted below 1% *in vivo*. Additionally, failed HDR does not preclude a successful double-stranded break at the target site, which could produce unintended insertions and deletions in *HBB* and potentially loss-of-function β-globin alleles (Canver et al., 2014). A more attractive strategy is disrupting the erythroid-specific enhancer element of *BCL11A* (Bauer et al., 2013; Canver et al., 2015; Xu et al., 2011). In either of these therapies, HSPCs will be isolated from subjects, edited *ex vivo*, and then returned to the same individuals via autologous bone marrow transplantation. Advantages of targeting the enhancer are numerous: 1) NHEJ-induced indels are likely to disrupt the enhancer function; 2) human life is tolerant to genetic variation in the enhancer (as evidenced by GWAS); and 3) mutations potently induce HbF.

Our experiments show that CRISPR/Cas9-mediated disruption of the *HBG* promoter CCAAT box induces HbF to similar levels achieved by genome editing the *BCL11A* enhancer. Perhaps the greatest strength of our approach is that we recreate a mutation that is carried by normal, healthy individuals with HPFH. With any of the proposed genome editing strategies for hemoglobinopathies, future efforts will focus on optimizing delivery and on-target editing in long-term repopulating stem cells. Whereas the majority of our experiments utilized lentiviral delivery of gRNA/Cas9, genome editing for therapy would need to be transient to avoid the continued, lifelong genome editing in hematopoietic stem cells. Currently, transient expression techniques for suspension cells require electroporation (DNA, RNA, or protein) or an integration-defective lentivirus (Kim and Kim, 2014). Recent work suggested by electroporation of Cas9 complexes pre-loaded with gRNA preserves the viability of $CD34^+$ cells, but studies are necessary to demonstrate that a sub-population of HSCs are edited and can repopulate bone marrow niche (Hendel et al., 2015). A promising, yet untested approach for any suspension cells or HSPCs is a protein transduction method that utilizes a combination of NaCl hypertonicity-induced macropinocytosis and propanebetaine to allow cellular uptake of proteins from the environment (D'Astolfo et al., 2015).

Furthermore, it is not yet known which set of genome editing tools would be most efficient to target -102 to -114 of the γ-globin promoters. Zinc finger nucleases and TALENs (Transcription Activator-Like Effector Nucleases) should also be tested in parallel to directly compare the on-target editing rates in stem cells. These efficiencies

would be evaluated in an immunodeficient xenograft mouse model (NOD.Cg-*Prkd*^{SCID}Il2rg^{tm1Wjil}/SzJ (NSG)), which allows the engraftment of human HSCs. Xenograft mouse models are suboptimal for functional assessments and fetal hemoglobin production because human-derived erythroblasts fail to develop in these mice. However, evaluation of the editing rates in human cells after 12 weeks will demonstrate the editing efficiency in long-term repopulating HSCs.

Off-target effects

Genome editing approaches remain untested in patients and numerous challenges must be overcome for safe and effective clinical translation. The unpredictability and error-prone nature of NHEJ-mediated repair poses many safety concerns. Off-target editing is also a particular concern in CRISPR/Cas9 targeting of this locus because the CCAAT box comprises 5 of the 20 nt in the gRNA sequence. Of note, about 30% of randomly surveyed promoters contained a CCAAT motif (Bucher, 1990). While we performed amplicon sequencing at off-target sites across the genome, as predicted by bioinformatics approaches, this method is sub-optimal in sensitivity for the detection of rare mutation events and events outside our candidate loci. An alternative method, Cas9 ChIP-sequencing, reveals the nature of Cas9 binding genome-wide, but Cas9 may simply bind and not create a double-strand break. We plan to pursue recently developed genome-wide off-target analysis called GUIDE-sequencing, where a short oligonucleotide tag integrates at sites of DNA breaks at the time of editing (Tsai et al., 2015). PCR and

sequencing of the tagged regions allow identification of the loci surrounding the DNA break. These described approaches require the sensitivity to identify rare editing events potentially anywhere in the human genome, and sequencing multiple genomes as in a bulk pool of cells limits the sensitivity of these approaches. Therefore, we also plan to genome edit -102 to -114 of γ-globin promoters in a humanized mouse model of SCD. We will attempt editing hematopoietic stem cells with subsequent bone marrow transplantation, in addition to germline editing of zygotes. In addition to assessing the extent of SCD phenotypic rescue with editing, we will also evaluate mice for hematopoietic neoplasms.

Extension to non-editing approaches to gene therapy

Gene therapy strategies for SCD and β thalassemia are currently in clinical trials for β-globin gene replacement. In this approach, subjects undergo autologous bone marrow transplantation with cells that are transduced with lentivirus encoding an anti-sickling form of β-globin (β^{T87Q}). Initial results with LentiGlobin HPV569 and LentiGlobin BB305 are highly promising – the first treated subject became and has remained transfusion independent, now 7 years after treatment (Cavazzana-Calvo et al., 2010). However, the following subjects with more severe forms of β thalassemia and SCD had more varying responses to therapy. Lentiviral gene therapies to exogenously express γ-globin are also being developed (Perumbeti et al., 2009; Pestina et al., 2009; 2015; Wilber et al., 2011). The commonalities amongst these vectors are the enhancer

and promoter elements driving globin expression, including a mini-LCR (HS2, HS3, and HS4) followed by the β-globin promoter. We posit that higher exogenous levels of γ-globin could be achieved by utilizing the γ-globin promoter adapted with the 13-nt HPFH deletion.

6.2.3 How does the CCAAT box repress HbF expression?

Ongoing and future studies are now focused on the mechanisms of HbF regulation mediated by the regulatory sequence in the γ−globin promoter. In a clonal analysis of edited HUDEP-2 cells, we showed that 1- to 4-nt deletions within the CCAAT box is sufficient for high HbF expression. The -114 γ-globin promoter CCAAT box likely mediates postnatal γ-to-β globin switching by recruiting transcriptional repressors in a developmentally regulated fashion (Forget, 1998; Stamatoyannopoulos, 2005). Candidate CCAAT box binding proteins include COUP-TFII (NR2F2, NF-E3 (Mantovani et al., 1989; Ronchi et al., 1995))(Liberati et al., 2001), NF-Y (CP-1, CBF) (Liberati et al., 2001; Zhu et al., 2012), CDP (Mantovani et al., 1989; Superti-Furga et al., 1988), and C/EBP (Superti-Furga et al., 1988). An overlapping motif binds nuclear hormone receptors TR2 and TR4 (Tanabe et al., 2002). Although the molecular triggers of γ-to-β globin switching are not fully defined, we plan to perform chromatin-immunoprecipitation (ChIP) experiments to determine the occupancy of these candidate DNA-binding proteins. We will also test whether *BCL11A* binding is altered by CCAAT

box mutations by ChIP, despite the absence of a consensus sequence or previously described ChIP signal.

Our studies suggest that altering the γ-globin promoter CCAAT box reverses the γ–to-β switch. Instead of simply inducing γ-globin, we show that absolute levels of β-globin are concomitantly reduced. This suggests that the edited promoter region adjusts the long-range contacts with LCR by promoting enhancer-promoter interactions with *HBG1* and *HBG2*, as opposed to *HBB*. We plan to test the nature of these interactions in edited, high-HbF and unedited HUDEP-2 cells by performing 3C (chromosome conformation capture)-based chromosomal interaction studies. In an extended 3C method called Capture C, we will cross-link DNA-bound proteins in cells, restriction enzyme digest the genome, and ligate fragments that are in spatial proximity to one another. We will build sequencing libraries and enrich for sequences interacting with the LCR by performing a nucleotide capture step with biotinylated oligonucleotide probes (Hughes et al., 2013; 2014).

Deng et al. have shown that forcing the chromatin to form a loop between the LCR and the γ-globin promoter raised HbF and commensurately reduced β-globin expression (Deng et al., 2014). Interestingly, the zinc finger domains used in this study target a region 1 nt upstream of the 13-nt HPFH deletion (Wilber et al., 2011). It is possible that the CCAAT box is merely part of the nucleotide sequence required γ-globin

repression, and that mutagenesis of the CCAAT box sufficiently removes the critical DNA-binding sites for either looping factors or transcriptional repressors.

6.2.4 Identifying additional cis-regulatory elements

We isolated several HbF-high HUDEP-2 clones in which simultaneous Cas9-gRNA-induced DNA breaks in both promoters of *HBG1* and *HBG2* caused deletion of the intervening 5.2-kb sequence. While the occurrence of the intervening deletion was predicted based on gRNA recognition of the duplicated promoters regions, the high-HbF phenotypes of these clones was a surprising finding. The 5.2-kb deletion generated removes the entire *HBG2* gene. One hypothesis to explain the effects of the large deletion de-represses γ-globin is the removal of an additional negative regulatory element lying within *HBG2* (intronic) or between (intergenic) *HBG1* and *HBG2*. We plan to investigate this further by generating and analyzing HUDEP-2 clones with a series of deletions in the region to identify critical sequences. Identifying and studying novel regulatory sequences will elucidate new mechanisms of γ-globin regulation and potential therapeutic targets.

6.3 Concluding remarks

In this thesis, we have leveraged insights from human genetic variation to perform focused studies on two different aspects of erythroid biology. In Chapters 3 and 4, we performed follow-up experiments to test the hypothesis that GWAS inspired – that SNPs

in linkage disequilibrium with *TRIM58* associated with deviations in RBC parameters implicated the gene in erythroid development. In Chapter 6, we used genome-editing techniques to recreate a rare genetic variant associated with HPFH, a condition with an oppositely extreme effect size. While seemingly different at the surface, studies of both Trim58 and γ-globin gene regulation also highlight the intricacies of the dynamic expression (both at the level of transcription and protein) changes that occur during terminal erythropoiesis. Finally, our findings provide an example of how knowledge learned from the study of patients can provide approaches for rationally designed therapies.

Chapter 7 BIBLIOGRAPHY

Akinsheye, I., Alsultan, A., Solovieff, N., Ngo, D., Baldwin, C.T., Sebastiani, P., Chui, D.H.K., and Steinberg, M.H. (2011). Fetal hemoglobin in sickle cell anemia. Blood *118*, 19–27.

Alami, R., Bender, M.A., Feng, Y.Q., Fiering, S.N., Hug, B.A., Ley, T.J., Groudine, M., and Bouhassira, E.E. (2000). Deletions within the mouse beta-globin locus control region preferentially reduce beta(min) globin gene expression. Genomics *63*, 417–424.

An, X., Schulz, V.P., Li, J., Wu, K., Liu, J., Xue, F., Hu, J., Mohandas, N., and Gallagher, P.G. (2014). Global transcriptome analyses of human and murine terminal erythroid differentiation. Blood *123*, 3466–3477.

Basak, A., Hancarova, M., Ulirsch, J.C., Balci, T.B., Trkova, M., Pelisek, M., Vlckova, M., Muzikova, K., Cermak, J., Trka, J., et al. (2015). BCL11A deletions result in fetal hemoglobin persistence and neurodevelopmental alterations. J. Clin. Invest. *125*, 2363–2368.

Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L., et al. (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science *342*, 253–257.

Behe, M.J., and Englander, S.W. (1979). Mixed gelation theory. Kinetics, equilibrium and gel incorporation in sickle hemoglobin mixtures. J. Mol. Biol. *133*, 137–160.

Bender, M.A., Mehaffey, M.G., Telling, A., Hug, B., Ley, T.J., Groudine, M., and Fiering, S. (2000). Independent formation of DnaseI hypersensitive sites in the murine beta-globin locus control region. Blood *95*, 3600–3604.

Bender, M., Thon, J.N., Ehrlicher, A.J., Wu, S., Mazutis, L., Deschmann, E., Sola-Visner, M., Italiano, J.E., and Hartwig, J.H. (2015). Microtubule sliding drives proplatelet elongation and is dependent on cytoplasmic dynein. Blood *125*, 860–868.

Berry, M., Grosveld, F., and Dillon, N. (1992). A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. Nature *358*, 499–502.

Biris, N., Yang, Y., Taylor, A.B., Tomashevski, A., Guo, M., Hart, P.J., Diaz-Griffero, F., and Ivanov, D.N. (2012). Structure of the rhesus monkey TRIM5α PRYSPRY domain, the HIV capsid recognition module. Proceedings of the National Academy of Sciences *109*, 13278–13283.

Blanc, L., Liu, J., Vidal, M., Chasis, J.A., An, X., and Mohandas, N. (2009). The water channel aquaporin-1 partitions into exosomes during reticulocyte maturation: implication

for the regulation of cell volume. Blood *114*, 3928–3934.

Blaybel, R., Théoleyre, O., Douablin, A., and Baklouti, F. (2008). Downregulation of the Spi-1/PU.1 oncogene induces the expression of TRIM10/HERF1, a key factor required for terminal erythroid cell differentiation and survival. Cell Res. *18*, 834–845.

Bucher, P. (1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J. Mol. Biol. *212*, 563–578.

Campbell, A.E., Wilkinson-White, L., Mackay, J.P., Matthews, J.M., and Blobel, G.A. (2013). Analysis of disease-causing GATA1 mutations in murine gene complementation systems. Blood *121*, 5218–5227.

Canver, M.C., Bauer, D.E., Dass, A., Yien, Y.Y., Chung, J., Masuda, T., Maeda, T., Paw, B.H., and Orkin, S.H. (2014). Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. J Biol Chem *289*, 21312–21324.

Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature *527*, 192–197.

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.-F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat. Genet. *32*, 623–626.

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., et al. (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature *467*, 318–322.

Chasis, J.A., Prenant, M., Leung, A., and Mohandas, N. (1989). Membrane assembly and remodeling during reticulocyte maturation. Blood *74*, 1112–1120.

Chassanidis, C., Kalamaras, A., Phylactides, M., Pourfarzad, F., Likousi, S., Maroulis, V., Papadakis, M.N., Vamvakopoulos, N.K., Aleporou-Marinou, V., Patrinos, G.P., et al. (2009). The Hellenic type of nondeletional hereditary persistence of fetal hemoglobin results from a novel mutation (g.-109G>T) in the HBG2 gene promoter. Ann. Hematol. *88*, 549–555.

Cheloufi, S., Santos, dos, C.O., Chong, M.M.W., and Hannon, G.J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature *465*, 584–589.

Chen, C.Y., Pajak, L., Tamburlin, J., Bofinger, D., and Koury, S.T. (2002). The effect of

proteasome inhibitors on mammalian erythroid terminal differentiation. Exp. Hematol. *30*, 634–639.

Chen, K., Liu, J., Heck, S., Chasis, J.A., An, X., and Mohandas, N. (2009). Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis. Proceedings of the National Academy of Sciences *106*, 17413–17418.

Cheng, Y., Wu, W., Kumar, S.A., Yu, D., Deng, W., Tripic, T., King, D.C., Chen, K.-B., Zhang, Y., Drautz, D., et al. (2009). Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. Genome Res. *19*, 2172–2184.

Ciechanover, A., Hod, Y., and Hershko, A. (2012). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. 1978.

Cifuentes, D., Xue, H., Taylor, D.W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G.J., Lawson, N.D., et al. (2010). A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. Science *328*, 1694–1698.

Collins, F.S., Metherall, J.E., Yamakawa, M., Pan, J., Weissman, S.M., and Forget, B.G. (1985). A point mutation in the A gamma-globin gene promoter in Greek hereditary persistence of fetal haemoglobin. Nature *313*, 325–326.

D'Astolfo, D.S., Pagliero, R.J., Pras, A., Karthaus, W.R., Clevers, H., Prasad, V., Lebbink, R.J., Rehmann, H., and Geijsen, N. (2015). Efficient intracellular delivery of native proteins. Cell *161*, 674–690.

D'Cruz, A.A., Kershaw, N.J., Chiang, J.J., Wang, M.K., Nicola, N.A., Babon, J.J., Gack, M.U., and Nicholson, S.E. (2013). Crystal Structure of the TRIM25 B30.2 (PRYSPRY) domain: A Key Component of Antiviral Signaling. Biochem. J.

de Vasconcellos, J.F., Fasano, R.M., Lee, Y.T., Kaushal, M., Byrnes, C., Meier, E.R., Anderson, M., Rabel, A., Braylan, R., Stroncek, D.F., et al. (2014). LIN28A expression reduces sickling of cultured human erythrocytes. PLoS ONE *9*, e106924.

Deng, W., Rupon, J.W., Krivega, I., Breda, L., Motta, I., Jahn, K.S., Reik, A., Gregory, P.D., Rivella, S., Dean, A., et al. (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell *158*, 849–860.

Deshaies, R.J., and Joazeiro, C.A.P. (2009). RING Domain E3 Ubiquitin Ligases. Http://Proxy.Library.Upenn.Edu:2146/10.1146/Annurev.Biochem.78.101807.093809 *78*, 399–434.

Diagouraga, B., Grichine, A., Fertin, A., Wang, J., Khochbin, S., and Sadoul, K. (2014). Motor-driven marginal band coiling promotes cell shape change during platelet activation. J Cell Biol *204*, 177–185.

Dolznig, H., Bartunek, P., Nasmyth, K., Müllner, E.W., and Beug, H. (1995). Terminal differentiation of normal chicken erythroid progenitors: shortening of G1 correlates with loss of D-cyclin/cdk4 expression and altered cell size control. Cell Growth Differ. *6*, 1341–1352.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

E, S., and D, D. (1967). An electron microscopic study of nuclear elimination from the late erythroblast. Journal of Cell Biology 1–11.

Egan, E.S., Jiang, R.H.Y., Moechtar, M.A., Barteneva, N.S., Weekes, M.P., Nobre, L.V., Gygi, S.P., Paulo, J.A., Frantzreb, C., Tani, Y., et al. (2015). Malaria. A forward genetic screen identifies erythrocyte CD55 as essential for Plasmodium falciparum invasion. Science *348*, 711–714.

Erpapazoglou, Z., Dhaoui, M., Pantazopoulou, M., Giordano, F., Mari, M., Leon, S., Raposo, G., Reggiori, F., and Haguenauer-Tsapis, R. (2012). A dual role for K63-linked ubiquitin chains in multivesicular body biogenesis and cargo sorting. Molecular Biology of the Cell *23*, 2170–2183.

Etlinger, J.D., and Goldberg, A.L. (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. Proc Natl Acad Sci USA *74*, 54–58.

Fabrini, R., De Luca, A., Stella, L., Mei, G., Orioni, B., Ciccone, S., Federici, G., Bello, Lo, M., and Ricci, G. (2009). Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. Biochemistry *48*, 10473–10482.

Feingold, E.A., and Forget, B.G. (1989). The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin occurs within an erythroid DNA domain remote from the beta-globin gene cluster. Blood *74*, 2178–2186.

Forget, B.G. (1998). Molecular basis of hereditary persistence of fetal hemoglobin. Ann. N. Y. Acad. Sci. *850*, 38–44.

Forrester, W.C., Thompson, C., Elder, J.T., and Groudine, M. (1986). A developmentally stable chromatin structure in the human beta-globin gene cluster. Proc Natl Acad Sci USA *83*, 1359–1363.

Freemont, P.S., Hanson, I.M., and Trowsdale, J. (1991). A novel cysteine-rich sequence motif. Cell *64*, 483–484.

Fucharoen, S., Shimizu, K., and Fukumaki, Y. (1990). A novel C-T transition within the distal CCAAT motif of the G gamma-globin gene in the Japanese HPFH: implication of factor binding in elevated fetal globin expression. Nucleic Acids Res. *18*, 5245–5253.

Funnell, A.P.W., Prontera, P., Ottaviani, V., Piccione, M., Giambona, A., Maggio, A., Ciaffoni, F., Stehling-Sun, S., Marra, M., Masiello, F., et al. (2015). 2p15-p16.1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment. Blood *126*, 89–93.

Ganesh, S.K., Zakai, N.A., van Rooij, F.J.A., Soranzo, N., Smith, A.V., Nalls, M.A., Chen, M.-H., Kottgen, A., Glazer, N.L., Dehghan, A., et al. (2009). Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. Nat. Genet. *41*, 1191–1198.

Giani, F.C., Fiorini, C., Wakabayashi, A., Ludwig, L.S., Salem, R.M., Jobaliya, C.D., Regan, S.N., Ulirsch, J.C., Liang, G., Steinberg-Shemer, O., et al. (2015). Targeted Application of Human Genetic Variation Can Improve Red Blood Cell Production from Stem Cells. Cell Stem Cell.

Gieger, C., Radhakrishnan, A., Cvejic, A., Tang, W., Porcu, E., Pistis, G., Serbanovic-Canic, J., Elling, U., Goodall, A.H., Labrune, Y., et al. (2011). New gene functions in megakaryopoiesis and platelet formation. Nature *480*, 201–208.

Gifford, S.C., Derganc, J., Shevkoplyas, S.S., Yoshida, T., and Bitensky, M.W. (2006). A detailed study of time-dependent changes in human red blood cells: from reticulocyte maturation to erythrocyte senescence. Br. J. Haematol. *135*, 395–404.

Gilman, J.G., Mishima, N., Wen, X.J., Stoming, T.A., Lobel, J., and Huisman, T.H. (1988). Distal CCAAT box deletion in the A gamma globin gene of two black adolescents with elevated fetal A gamma globin. Nucleic Acids Res. *16*, 10635–10642.

Goldstone, D.C., Walker, P.A., Calder, L.J., Coombs, P.J., Kirkpatrick, J., Ball, N.J., Hilditch, L., Yap, M.W., Rosenthal, P.B., Stoye, J.P., et al. (2014). Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. Proceedings of the National Academy of Sciences *111*, 9609–9614.

Goodman, S.R., Kurdia, A., Ammann, L., Kakhniashvili, D., and Daescu, O. (2007). The human red blood cell proteome and interactome. Exp. Biol. Med. (Maywood) *232*, 1391–1408.

Grebien, F., Dolznig, H., Beug, H., and Mullner, E.W. (2005). Cell size control: new

evidence for a general mechanism. Cell Cycle *4*, 418–421.

Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F., and Kay, M.A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature *441*, 537–541.

Grütter, C., Briand, C., Capitani, G., Mittl, P.R.E., Papin, S., Tschopp, J., and Grütter, M.G. (2006). Structure of the PRYSPRY-domain: implications for autoinflammatory diseases. FEBS Lett. *580*, 99–106.

Guda, S., Brendel, C., Renella, R., Du, P., Bauer, D.E., Canver, M.C., Grenier, J.K., Grimson, A.W., Kamran, S.C., Thornton, J., et al. (2015). miRNA-embedded shRNAs for Lineage-specific BCL11A Knockdown and Hemoglobin F Induction. Mol. Ther. *23*, 1465–1474.

Harada, H., Harada, Y., O'Brien, D.P., Rice, D.S., Naeve, C.W., and Downing, J.R. (1999). HERF1, a novel hematopoiesis-specific RING finger protein, is required for terminal differentiation of erythroid cells. Mol Cell Biol *19*, 3808–3815.

Harmston, N., and Lenhard, B. (2013). Chromatin and epigenetic features of long-range gene regulation. Nucleic Acids Res. *41*, 7185–7199.

Hattangadi, S.M., Martinez-Morilla, S., Patterson, H.C., Shi, J., Burke, K., Avila-Figueroa, A., Venkatesan, S., Wang, J., Paulsen, K., Görlich, D., et al. (2014). Histones to the cytosol: exportin 7 is essential for normal terminal erythroid nuclear maturation. Blood *124*, 1931–1940.

Hemann, M.T., Fridman, J.S., Zilfou, J.T., Hernando, E., Paddison, P.J., Cordon-Cardo, C., Hannon, G.J., and Lowe, S.W. (2003). An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. Nat. Genet. *33*, 396–400.

Hendel, A., Bak, R.O., Clark, J.T., Kennedy, A.B., Ryan, D.E., Roy, S., Steinfeld, I., Lunstad, B.D., Kaiser, R.J., Wilkens, A.B., et al. (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. Nat. Biotechnol. *33*, 985–989.

Hendricks, A.G., Lazarus, J.E., Perlson, E., Gardner, M.K., Odde, D.J., Goldman, Y.E., and Holzbaur, E.L.F. (2012). Dynein tethers and stabilizes dynamic microtubule plus ends. Curr Biol *22*, 632–637.

Henthorn, P.S., Smithies, O., and Mager, D.L. (1990). Molecular analysis of deletions in the human beta-globin gene cluster: deletion junctions and locations of breakpoints. Genomics *6*, 226–237.

Herrick, Irons (1910). Peculiar Elongated and Sickle-shaped red blood corpuscles in a case of severe anemia (Archives of Internal Medicine).

Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J Biol Chem *258*, 8206–8214.

Hoban, M.D., Cost, G.J., Mendel, M.C., Romero, Z., Kaufman, M.L., Joglekar, A.V., Ho, M., Lumaquin, D., Gray, D., Lill, G.R., et al. (2015). Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. Blood *125*, 2597–2604.

Hoban, M.D., Orkin, S.H., and Bauer, D.E. (2016). Genetic treatment of a molecular disorder: gene therapy approaches to sickle cell disease. Blood *127*, 839–848.

Howell, B.J., McEwen, B.F., Canman, J.C., Hoffman, D.B., Farrar, E.M., Rieder, C.L., and Salmon, E.D. (2001). Cytoplasmic dynein/dynactin drives kinetochore protein transport to the spindle poles and has a role in mitotic spindle checkpoint inactivation. Journal of Cell Biology *155*, 1159–1172.

Hughes, J.R., Lower, K.M., Dunham, I., Taylor, S., De Gobbi, M., Sloane-Stanley, J.A., McGowan, S., Ragoussis, J., Vernimmen, D., Gibbons, R.J., et al. (2013). High-resolution analysis of cis-acting regulatory networks at the α-globin locus. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *368*, 20120361.

Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., and Higgs, D.R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat. Genet. *46*, 205–212.

Huisman, T.H., Schroeder, W.A., Charache, S., Bethlenfalvay, N.C., Bouver, N., Shelton, J.R., Shelton, J.B., and Apell, G. (1971). Hereditary persistence of fetal hemoglobin. Heterogeneity of fetal hemoglobin in homozygotes and in conjunction with -thalassemia. N. Engl. J. Med. *285*, 711–716.

Jackson, A.L., and Linsley, P.S. (2004). Noise amidst the silence: off-target effects of siRNAs? Trends Genet. *20*, 521–524.

Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L., and Linsley, P.S. (2006). Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. Rna *12*, 1179–1187.

Jagadeeswaran, P., Tuan, D., Forget, B.G., and Weissman, S.M. (1982). A gene deletion ending at the midpoint of a repetitive DNA sequence in one form of hereditary persistence of fetal haemoglobin. Nature *296*, 469–470.

James, L.C., Keeble, A.H., Khan, Z., Rhodes, D.A., and Trowsdale, J. (2007). Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. Proc Natl Acad Sci USA *104*, 6200–6205.

Jayapal, S.R., Lee, K.L., Ji, P., Kaldis, P., Lim, B., and Lodish, H.F. (2010). Down-regulation of Myc is essential for terminal erythroid maturation. J Biol Chem *285*, 40252–40265.

Ji, P., Jayapal, S.R., and Lodish, H.F. (2008). Enucleation of cultured mouse fetal erythroblasts requires Rac GTPases and mDia2. Nat Cell Biol *10*, 314–321.

Ji, P., Yeh, V., Ramirez, T., Murata-Hori, M., and Lodish, H.F. (2010). Histone deacetylase 2 is required for chromatin condensation and subsequent enucleation of cultured mouse fetal erythroblasts. Haematologica *95*, 2013–2021.

Joly, P., Lacan, P., Garcia, C., Couprie, N., and Francina, A. (2009). Identification and molecular characterization of four new large deletions in the beta-globin gene cluster. Blood Cells Mol. Dis. *43*, 53–57.

Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., Nakamura, Y., and Kamatani, N. (2010). Genome-wide association study of hematological and biochemical traits in a Japanese population. Nat. Genet. *42*, 210–215.

Keerthivasan, G., Small, S., Liu, H., Wickrema, A., and Crispino, J.D. (2010). Vesicle trafficking plays a novel role in erythroblast enucleation. Blood *116*, 3331–3340.

Keerthivasan, G., Wickrema, A., and Crispino, J.D. (2011). Erythroblast Enucleation. Stem Cells Int *2011*, 1–9.

Khandros, E., and Weiss, M.J. (2010). Protein quality control during erythropoiesis and hemoglobin synthesis. Hematol. Oncol. Clin. North Am. *24*, 1071–1088.

Khoriaty, R., Vasievich, M.P., Jones, M., Everett, L., Chase, J., Tao, J., Siemieniak, D., Zhang, B., Maillard, I., and Ginsburg, D. (2014). Absence of a red blood cell phenotype in mice with hematopoietic deficiency of SEC23B. Mol Cell Biol *34*, 3721–3734.

Kidoguchi, K., Ogawa, M., Karam, J.D., and Martin, A.G. (1978). Augmentation of fetal hemoglobin (HbF) synthesis in culture by human erythropoietic precursors in the marrow and peripheral blood: studies in sickle cell anemia and nonhemoglobinopathic adults. Blood *52*, 1115–1124.

Kim, H., and Kim, J.-S. (2014). A guide to genome engineering with programmable nucleases. Nat. Rev. Genet. *15*, 321–334.

Kingsley, P.D., Greenfest-Allen, E., Frame, J.M., Bushnell, T.P., Malik, J., McGrath, K.E., Stoeckert, C.J., and Palis, J. (2013). Ontogeny of erythroid gene expression. Blood *121*, e5–e13.

Kobayashi, I., Ubukawa, K., Sugawara, K., Asanuma, K., Guo, Y.-M., Yamashita, J., Takahashi, N., Sawada, K., and Nunomura, W. (2016). Erythroblast enucleation is a dynein-dependent process. Exp. Hematol. *44*, 247–256.e12.

Koliopoulos, M.G., Esposito, D., Christodoulou, E., Taylor, I.A., and Rittinger, K. (2016). Functional role of TRIM E3 ligase oligomerization and regulation of catalytic activity. The EMBO Journal.

Kollias, G., Wrighton, N., Hurst, J., and Grosveld, F. (1986). Regulated expression of human A gamma-, beta-, and hybrid gamma beta-globin genes in transgenic mice: manipulation of the developmental expression patterns. Cell *46*, 89–94.

Konstantinidis, D.G., Pushkaran, S., Johnson, J.F., Cancelas, J.A., Manganaris, S., Harris, C.E., Williams, D.A., Zheng, Y., and Kalfa, T.A. (2012). Signaling and cytoskeletal requirements in erythroblast enucleation. Blood *119*, 6118–6127.

Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., Tani, K., and Nakamura, Y. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. PLoS ONE *8*, e59890.

Lekomtsev, S., Su, K.-C., Pye, V.E., Blight, K., Sundaramoorthy, S., Takaki, T., Collinson, L.M., Cherepanov, P., Divecha, N., and Petronczki, M. (2012). Centralspindlin links the mitotic spindle to the plasma membrane during cytokinesis. Nature *492*, 276–279.

Lettre, G., Sankaran, V.G., Bezerra, M.A.C., Araújo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N., et al. (2008). DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proceedings of the National Academy of Sciences *105*, 11869–11874.

Li, Q., Peterson, K.R., Fang, X., and Stamatoyannopoulos, G. (2002). Locus control regions. Blood *100*, 3077–3086.

Li, W., Bengtson, M.H., Ulbrich, A., Matsuda, A., Reddy, V.A., Orth, A., Chanda, S.K., Batalov, S., and Joazeiro, C.A.P. (2008). Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. PLoS ONE *3*, e1487.

Li, Y., Wu, H., Wu, W., Zhuo, W., Liu, W., Zhang, Y., Cheng, M., Chen, Y.-G., Gao, N.,

Yu, H., et al. (2014). Structural insights into the TRIM family of ubiquitin E3 ligases. Cell Res. *24*, 762–765.

Liang, R., Campreciós, G., Kou, Y., McGrath, K., Nowak, R., Catherman, S., Bigarella, C.L., Rimmelé, P., Zhang, X., Gnanapragasam, M.N., et al. (2015). A Systems Approach Identifies Essential FOXO3 Functions at Key Steps of Terminal Erythropoiesis. PLoS Genet. *11*, e1005526.

Liberati, C., Cera, M.R., Secco, P., Santoro, C., Mantovani, R., Ottolenghi, S., and Ronchi, A. (2001). Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. J Biol Chem *276*, 41700–41709.

Liljeholm, M., Irvine, A.F., Vikberg, A.-L., Norberg, A., Month, S., Sandström, H., Wahlin, A., Mishima, M., and Golovleva, I. (2013). Congenital dyserythropoietic anemia type III (CDA III) is caused by a mutation in kinesin family member, KIF23. Blood.

Liu, J., Guo, X., Mohandas, N., Chasis, J.A., and An, X. (2010). Membrane remodeling during reticulocyte maturation. Blood *115*, 2021–2027.

Ludwig, L.S., Cho, H., Wakabayashi, A., Eng, J.C., Ulirsch, J.C., Fleming, M.D., Lodish, H.F., and Sankaran, V.G. (2015). Genome-wide association study follow-up identifies cyclin A2 as a regulator of the transition through cytokinesis during terminal erythropoiesis. Am. J. Hematol. *90*, 386–391.

Lux, C.T., Yoshimoto, M., McGrath, K., Conway, S.J., Palis, J., and Yoder, M.C. (2008). All primitive and definitive hematopoietic progenitor cells emerging before E10 in the mouse embryo are products of the yolk sac. Blood *111*, 3435–3438.

Ma, S., Triviños-Lagos, L., Gräf, R., and Chisholm, R.L. (1999). Dynein intermediate chain mediated dynein-dynactin interaction is required for interphase microtubule organization and centrosome replication and separation in Dictyostelium. Journal of Cell Biology *147*, 1261–1274.

Mantovani, R., Superti-Furga, G., Gilman, J., and Ottolenghi, S. (1989). The deletion of the distal CCAAT box region of the A gamma-globin gene in black HPFH abolishes the binding of the erythroid specific protein NFE3 and of the CCAAT displacement protein. Nucleic Acids Res. *17*, 6681–6691.

Marín, I. (2012). Origin and diversification of TRIM ubiquitin ligases. PLoS ONE *7*, e50030.

McGrath, K., and Palis, J. (2008). Ontogeny of erythropoiesis in the mammalian embryo. Curr. Top. Dev. Biol. *82*, 1–22.

McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Rep *11*, 1892–1904.

McGrath, K.E., Frame, J.M., Fromm, G.J., Koniski, A.D., Kingsley, P.D., Little, J., Bulger, M., and Palis, J. (2011). A transient definitive erythroid lineage with unique regulation of the β-globin locus in the mammalian embryo. Blood *117*, 4600–4608.

McGrath, K.E., Kingsley, P.D., Koniski, A.D., Porter, R.L., Bushnell, T.P., and Palis, J. (2008). Enucleation of primitive erythroid cells generates a transient population of "pyrenocytes" in the mammalian fetus. Blood *111*, 2409–2417.

McKenney, R.J., Vershinin, M., Kunwar, A., Vallee, R.B., and Gross, S.P. (2010). LIS1 and NudE induce a persistent dynein force-producing state. Cell *141*, 304–314.

McKenney, R.J., Weil, S.J., Scherer, J., and Vallee, R.B. (2011). Mutually exclusive cytoplasmic dynein regulation by NudE-Lis1 and dynactin. J Biol Chem *286*, 39615–39622.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H., et al. (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat. Genet. *39*, 1197–1199.

Merryweather-Clarke, A.T., Atzberger, A., Soneji, S., Gray, N., Clark, K., Waugh, C., McGowan, S.J., Taylor, S., Nandi, A.K., Wood, W.G., et al. (2011). Global gene expression analysis of human erythroid progenitors. Blood *117*, e96–e108.

Mishima, M., Kaitna, S., and Glotzer, M. (2002). Central spindle assembly and cytokinesis require a kinesin-like protein/RhoGAP complex with microtubule bundling activity. Developmental Cell *2*, 41–54.

Napolitano, L.M., and Meroni, G. (2012). TRIM family: Pleiotropy and diversification through homomultimer and heteromultimer formation. IUBMB Life *64*, 64–71.

Noy-Lotan, S., Dgany, O., Lahmi, R., Marcoux, N., Krasnov, T., Yissachar, N., Ginsberg, D., Motro, B., Resnitzky, P., Yaniv, I., et al. (2009). Codanin-1, the protein encoded by the gene mutated in congenital dyserythropoietic anemia type I (CDAN1), is cell cycle-regulated. Haematologica *94*, 629–637.

Ozato, K., Shin, D.-M., Chang, T.-H., and Morse, H.C. (2008). TRIM family proteins and their emerging roles in innate immunity. Nat. Rev. Immunol. *8*, 849–860.

Palis, J., McGrath, K.E., and Kingsley, P.D. (1995). Initiation of hematopoiesis and vasculogenesis in murine yolk sac explants. Blood *86*, 156–163.

Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. Development *126*, 5073–5084.

Palstra, R.J., de Laat, W., and Grosveld, F. (2008). Beta-globin regulation and long-range interactions. Adv. Genet. *61*, 107–142.

Papayannopoulou, T.H., Brice, M., and Stamatoyannopoulos, G. (1976). Stimulation of fetal hemoglobin synthesis in bone marrow cultures from adult individuals. Proc Natl Acad Sci USA *73*, 2033–2037.

Papayannopoulou, T., Brice, M., and Stamatoyannopoulos, G. (1977). Hemoglobin F synthesis in vitro: evidence for control at the level of primitive erythroid stem cells. Proc Natl Acad Sci USA *74*, 2923–2927.

Paralkar, V.R., Mishra, T., Luan, J., Yao, Y., Kossenkov, A.V., Anderson, S.M., Dunagin, M., Pimkin, M., Gore, M., Sun, D., et al. (2014). Lineage and species-specific long noncoding RNAs during erythro-megakaryocytic development. Blood *123*, 1927–1937.

Pasini, E.M., Kirkegaard, M., Mortensen, P., Lutz, H.U., Thomas, A.W., and Mann, M. (2006). In-depth analysis of the membrane and cytosolic proteome of red blood cells. Blood *108*, 791–801.

Pasini, E.M., Kirkegaard, M., Salerno, D., Mortensen, P., Mann, M., and Thomas, A.W. (2008). Deep coverage mouse red blood cell proteome: a first comparison with the human red blood cell. Mol. Cell Proteomics *7*, 1317–1330.

Patel, S.R., Richardson, J.L., Schulze, H., Kahle, E., Galjart, N., Drabek, K., Shivdasani, R.A., Hartwig, J.H., and Italiano, J.E. (2005). Differential roles of microtubule assembly and sliding in proplatelet formation by megakaryocytes. Blood *106*, 4076–4085.

Pauling, L., Itano, H.A., Singer, S.J., and Wells, I.C. (1949). Sickle Cell Anemia, a Molecular Disease. Science *110*, 543–548.

Pavicic-Kaltenbrunner, V., Mishima, M., and Glotzer, M. (2007). Cooperative assembly of CYK-4/MgcRacGAP and ZEN-4/MKLP1 to form the centralspindlin complex. Molecular Biology of the Cell *18*, 4992–5003.

Perumbeti, A., Higashimoto, T., Urbinati, F., Franco, R., Meiselman, H.J., Witte, D., and Malik, P. (2009). A novel human gamma-globin gene vector for genetic correction of

sickle cell anemia in a humanized sickle mouse model: critical determinants for successful correction. Blood *114*, 1174–1185.

Pestina, T.I., Hargrove, P.W., Jay, D., Gray, J.T., Boyd, K.M., and Persons, D.A. (2009). Correction of murine sickle cell disease using gamma-globin lentiviral vectors to mediate high-level expression of fetal hemoglobin. Mol. Ther. *17*, 245–252.

Pestina, T.I., Hargrove, P.W., Zhao, H., Mead, P.E., Smeltzer, M.P., Weiss, M.J., Wilber, A., and Persons, D.A. (2015). Amelioration of murine sickle cell disease by nonablative conditioning and γ-globin gene-corrected bone marrow cells. Mol Ther Methods Clin Dev *2*, 15045.

Pimkin, M., Kossenkov, A.V., Mishra, T., Morrissey, C.S., Wu, W., Keller, C.A., Blobel, G.A., Lee, D., Beer, M.A., Hardison, R.C., et al. (2014). Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythro-megakaryopoiesis. Genome Res. *24*, 1932–1944.

Pishesha, N., Thiru, P., Shi, J., Eng, J.C., Sankaran, V.G., and Lodish, H.F. (2014). Transcriptional divergence and conservation of human and mouse erythropoiesis. Proceedings of the National Academy of Sciences *111*, 4103–4108.

Platt, O.S., Brambilla, D.J., Rosse, W.F., Milner, P.F., Castro, O., Steinberg, M.H., and Klug, P.P. (1994). Mortality in sickle cell disease. Life expectancy and risk factors for early death. N. Engl. J. Med. *330*, 1639–1644.

Platt, O.S., Thorington, B.D., Brambilla, D.J., Milner, P.F., Rosse, W.F., Vichinsky, E., and Kinney, T.R. (1991). Pain in sickle cell disease. Rates and risk factors. N. Engl. J. Med. *325*, 11–16.

Popova, E.Y., Krauss, S.W., Short, S.A., Lee, G., Villalobos, J., Etzell, J., Koury, M.J., Ney, P.A., Chasis, J.A., and Grigoryev, S.A. (2009). Chromatin condensation in terminally differentiating mouse erythroblasts does not involve special architectural proteins but depends on histone deacetylation. Chromosome Res. *17*, 47–64.

Quintyne, N.J., Gill, S.R., Eckley, D.M., Crego, C.L., Compton, D.A., and Schroer, T.A. (1999). Dynactin is required for microtubule anchoring at centrosomes. Journal of Cell Biology *147*, 321–334.

Renella, R., Roberts, N.A., Brown, J.M., De Gobbi, M., Bird, L.E., Hassanali, T., Sharpe, J.A., Sloane-Stanley, J., Ferguson, D.J.P., Cordell, J., et al. (2011). Codanin-1 mutations in congenital dyserythropoietic anemia type 1 affect HP1{alpha} localization in erythroblasts. Blood *117*, 6928–6938.

Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D.,

Zanaria, E., Messali, S., Cainarca, S., et al. (2001). The tripartite motif family identifies cell compartments. The EMBO Journal *20*, 2140–2151.

Richardson, J.L., Shivdasani, R.A., Boers, C., Hartwig, J.H., and Italiano, J.E. (2005). Mechanisms of organelle transport and capture along proplatelets during platelet production. Blood *106*, 4066–4075.

Ronchi, A.E., Bottardi, S., Mazzucchelli, C., Ottolenghi, S., and Santoro, C. (1995). Differential binding of the NFE3 and CP1/NFY transcription factors to the human gamma- and epsilon-globin CCAAT boxes. J Biol Chem *270*, 21934–21941.

Ronchi, A., Berry, M., Raguz, S., Imam, A., Yannoutsos, N., Ottolenghi, S., Grosveld, F., and Dillon, N. (1996). Role of the duplicated CCAAT box region in gamma-globin gene regulation and hereditary persistence of fetal haemoglobin. The EMBO Journal *15*, 143–149.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., and Stainier, D.Y.R. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature *524*, 230–233.

Roux-Dalvai, F., Gonzalez de Peredo, A., Simó, C., Guerrier, L., Bouyssié, D., Zanella, A., Citterio, A., Burlet-Schiltz, O., Boschetti, E., Righetti, P.G., et al. (2008). Extensive analysis of the cytoplasmic proteome of human erythrocytes using the peptide ligand library technology and advanced mass spectrometry. Mol. Cell Proteomics *7*, 2254–2269.

Rouzbeh, S., Kobari, L., Cambot, M., Mazurier, C., Hebert, N., Faussat, A.-M., Durand, C., Douay, L., and Lapillonne, H. (2015). Molecular signature of erythroblast enucleation in human embryonic stem cells. Stem Cells *33*, 2431–2441.

San Miguel, J., Bladé, J., Boccadoro, M., Cavenagh, J., Glasmacher, A., Jagannath, S., Lonial, S., Orlowski, R.Z., Sonneveld, P., and Ludwig, H. (2006). A practical update on the use of bortezomib in the management of multiple myeloma. Oncologist *11*, 51–61.

Sanchez, J.G., Okreglicka, K., Chandrasekaran, V., Welker, J.M., Sundquist, W.I., and Pornillos, O. (2014). The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer. Proceedings of the National Academy of Sciences *111*, 2494–2499.

Sankaran, V.G., and Weiss, M.J. (2015). Anemia: progress in molecular mechanisms and therapies. Nature Medicine *21*, 221–230.

Sankaran, V.G., Ludwig, L.S., Sicinska, E., Xu, J., Bauer, D.E., Eng, J.C., Patterson, H.C., Metcalf, R.A., Natkunam, Y., Orkin, S.H., et al. (2012). Cyclin D3 coordinates the cell cycle during differentiation to regulate erythrocyte size and number. Genes Dev *26*, 2075–2087.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K.A., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008a). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science *322*, 1839–1842.

Sankaran, V.G., Orkin, S.H., and Walkley, C.R. (2008b). Rb intrinsically promotes erythropoiesis by coupling cell cycle exit with mitochondrial biogenesis. Genes Dev *22*, 463–475.

Sankaran, V.G., Xu, J., Byron, R., Greisman, H.A., Fisher, C., Weatherall, D.J., Sabath, D.E., Groudine, M., Orkin, S.H., Premawardhena, A., et al. (2011). A functional element necessary for fetal hemoglobin silencing. N. Engl. J. Med. *365*, 807–814.

Schwarz, K., Iolascon, A., Verissimo, F., Trede, N.S., Horsley, W., Chen, W., Paw, B.H., Hopfner, K.-P., Holzmann, K., Russo, R., et al. (2009). Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. Nat. Genet. *41*, 936–940.

Shearstone, J.R., Pop, R., Bock, C., Boyle, P., Meissner, A., and Socolovsky, M. (2011). Global DNA demethylation during mouse erythropoiesis in vivo. Science *334*, 799–802.

Silver, L., and Palis, J. (1997). Initiation of murine embryonic erythropoiesis: a spatial analysis. Blood *89*, 1154–1164.

Simpson, C.F., and Kling, J.M. (1967). The mechanism of denucleation in circulating erythroblasts. Journal of Cell Biology *35*, 237–245.

Soni, S., Bala, S., Gwynn, B., Sahr, K.E., Peters, L.L., and Hanspal, M. (2006). Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. J Biol Chem *281*, 20181–20189.

Spence, J., Sadis, S., Haas, A.L., and Finley, D. (1995). A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. Mol Cell Biol *15*, 1265–1273.

Splinter, D., Tanenbaum, M.E., Lindqvist, A., Jaarsma, D., Flotho, A., Yu, K.L., Grigoriev, I., Engelsma, D., Haasdijk, E.D., Keijzer, N., et al. (2010). Bicaudal D2, dynein, and kinesin-1 associate with nuclear pore complexes and regulate centrosome and nuclear positioning during mitotic entry. PLoS Biol. *8*, e1000350.

ST, K., and MC, B. (1989). Cytoskeletal Distribution and Function during the maturation and enucleation of mammalian erythroblasts. Journal of Cell Biology 1–9.

Stamatoyannopoulos, G. (2005). Control of globin gene expression during development and erythroid differentiation. Exp. Hematol. *33*, 259–271.

Stonestrom, A.J., Hsu, S.C., Jahn, K.S., Huang, P., Keller, C.A., Giardine, B.M., Kadauke, S., Campbell, A.E., Evans, P., Hardison, R.C., et al. (2015). Functions of BET proteins in erythroid gene expression. Blood *125*, 2825–2834.

Streich, F.C., Ronchi, V.P., Connick, J.P., and Haas, A.L. (2013). Tripartite Motif Ligases Catalyze Polyubiquitin Chain Formation through a Cooperative Allosteric Mechanism. J Biol Chem *288*, 8209–8221.

Sunshine, H.R., Hofrichter, J., and Eaton, W.A. (1979). Gelation of sickle cell hemoglobin in mixtures with normal adult and fetal hemoglobins. J. Mol. Biol. *133*, 435–467.

Superti-Furga, G., Barberis, A., Schaffner, G., and Busslinger, M. (1988). The -117 mutation in Greek HPFH affects the binding of three nuclear factors to the CCAAT region of the gamma-globin gene. The EMBO Journal *7*, 3099–3107.

Tanabe, O., Katsuoka, F., Campbell, A.D., Song, W., Yamamoto, M., Tanimoto, K., and Engel, J.D. (2002). An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. The EMBO Journal *21*, 3434–3442.

Tanenbaum, M.E., Akhmanova, A., and Medema, R.H. (2011). Bi-directional transport of the nucleus by dynein and kinesin-1. Commun Integr Biol *4*, 21–25.

Thom, C.S., Traxler, E.A., Khandros, E., Nickas, J.M., Zhou, O.Y., Lazarus, J.E., Silva, A.P.G., Prabhu, D., Yao, Y., Aribeana, C., et al. (2014). Trim58 degrades Dynein and regulates terminal erythropoiesis. Developmental Cell *30*, 688–700.

Thrower, J.S., Hoffman, L., Rechsteiner, M., and Pickart, C.M. (2000). Recognition of the polyubiquitin proteolytic signal. The EMBO Journal *19*, 94–102.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Molecular Cell *10*, 1453–1465.

Truong, L.N., Li, Y., Shi, L.Z., Hwang, P.Y.-H., He, J., Wang, H., Razavian, N., Berns, M.W., and Wu, X. (2013). Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. Proceedings of the National Academy of Sciences *110*, 7720–7725.

Tsai, J.-W., Lian, W.-N., Kemal, S., Kriegstein, A.R., and Vallee, R.B. (2010). Kinesin 3 and cytoplasmic dynein mediate interkinetic nuclear migration in neural stem cells. Nat Neurosci *13*, 1463–1471.

Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens,

N., Khayter, C., Iafrate, A.J., Le, L.P., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. *33*, 187–197.

Tuan, D., Solomon, W., Li, Q., and London, I.M. (1985). The "beta-like-globin" gene domain in human erythroid cells. Proc Natl Acad Sci USA *82*, 6384–6388.

Ubukawa, K., Guo, Y.-M., Takahashi, M., Hirokawa, M., Michishita, Y., Nara, M., Tagawa, H., Takahashi, N., Komatsuda, A., Nunomura, W., et al. (2012). Enucleation of human erythroblasts involves non-muscle myosin IIB. Blood *119*, 1036–1044.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G., et al. (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proceedings of the National Academy of Sciences *105*, 1620–1625.

Ulirsch, J.C., Nandakumar, S.K., Wang, L., Giani, F.C., Zhang, X., Rogov, P., Melnikov, A., McDonel, P., Do, R., Mikkelsen, T.S., et al. (2016). Systematic Functional Dissection of Common Genetic Variation Affecting Red Blood Cell Traits. Cell *165*, 1530–1545.

Vallee, R.B., McKenney, R.J., and Ori-McKenney, K.M. (2012). Multiple modes of cytoplasmic dynein regulation. Nat Cell Biol *14*, 224–230.

van der Harst, P., Zhang, W., Mateo Leach, I., Rendon, A., Verweij, N., Sehmi, J., Paul, D.S., Elling, U., Allayee, H., Li, X., et al. (2012). Seventy-five genetic loci influencing the human red blood cell. Nature *492*, 369–375.

Versteeg, G.A., Rajsbaum, R., Sánchez-Aparicio, M.T., Maestre, A.M., Valdiviezo, J., Shi, M., Inn, K.-S., Fernandez-Sesma, A., Jung, J., and García-Sastre, A. (2013). The E3-ligase TRIM family of proteins regulates signaling pathways triggered by innate immune pattern-recognition receptors. Immunity *38*, 384–398.

Vierstra, J., Reik, A., Chang, K.-H., Stehling-Sun, S., Zhou, Y., Hinkley, S.J., Paschon, D.E., Zhang, L., Psatha, N., Bendana, Y.R., et al. (2015). Functional footprinting of regulatory DNA. Nature Publishing Group *12*, 927–930.

Wang, J., Ramirez, T., Ji, P., Jayapal, S.R., Lodish, H.F., and Murata-Hori, M. (2012). Mammalian erythroblast enucleation requires PI3K-dependent cell polarization. J Cell Sci *125*, 340–349.

Watanabe, S., De Zan, T., Ishizaki, T., Yasuda, S., Kamijo, H., Yamada, D., Aoki, T., Kiyonari, H., Kaneko, H., Shimizu, R., et al. (2013). Loss of a Rho-regulated actin nucleator, mDia2, impairs cytokinesis during mouse fetal erythropoiesis. Cell Rep *5*, 926–932.

Wefes, I., Mastrandrea, L.D., Haldeman, M., Koury, S.T., Tamburlin, J., Pickart, C.M., and Finley, D. (1995). Induction of ubiquitin-conjugating enzymes during terminal erythroid differentiation. Proc Natl Acad Sci USA *92*, 4982–4986.

Weinert, C., Morger, D., Djekic, A., Grütter, M.G., and Mittl, P.R.E. (2015). Crystal structure of TRIM20 C-terminal coiled-coil/B30.2 fragment: implications for the recognition of higher order oligomers. Sci Rep *5*, 10819.

Weiss, M.J., Yu, C., and Orkin, S.H. (1997). Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. Mol Cell Biol *17*, 1642–1651.

Welch, J.J., Watts, J.A., Vakoc, C.R., Yao, Y., Wang, H., Hardison, R.C., Blobel, G.A., Chodosh, L.A., and Weiss, M.J. (2004). Global regulation of erythroid gene expression by transcription factor GATA-1. Blood *104*, 3136–3147.

Wienert, B., Funnell, A.P.W., Norton, L.J., Pearson, R.C.M., Wilkinson-White, L.E., Lester, K., Vadolas, J., Porteus, M.H., Matthews, J.M., Quinlan, K.G.R., et al. (2015). Editing the genome to introduce a beneficial naturally occurring mutation associated with increased fetal globin. Nat Commun *6*, 7085.

Wilber, A., Hargrove, P.W., Kim, Y.-S., Riberdy, J.M., Sankaran, V.G., Papanikolaou, E., Georgomanoli, M., Anagnou, N.P., Orkin, S.H., Nienhuis, A.W., et al. (2011). Therapeutic levels of fetal hemoglobin in erythroid progeny of β-thalassemic CD34+ cells after lentiviral vector-mediated gene transfer. Blood *117*, 2817–2826.

Wilkinson, K.D., Urban, M.K., and Haas, A.L. (1980). Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. J Biol Chem *255*, 7529–7532.

Wilson, M.H., and Holzbaur, E.L.F. (2012). Opposing microtubule motors drive robust nuclear dynamics in developing muscle cells. J Cell Sci *125*, 4158–4169.

Wong, P.M., Chung, S.W., Reicheld, S.M., and Chui, D.H. (1986). Hemoglobin switching during murine embryonic development: evidence for two populations of embryonic erythropoietic progenitor cells. Blood *67*, 716–721.

Woo, J.-S., Suh, H.-Y., Park, S.-Y., and Oh, B.-H. (2006). Structural basis for protein recognition by B30.2/SPRY domains. Molecular Cell *24*, 967–976.

Wölwer, C.B., Pase, L.B., Pearson, H.B., Gödde, N.J., Lackovic, K., Huang, D.C.S., Russell, S.M., and Humbert, P.O. (2015). A Chemical Screening Approach to Identify Novel Key Mediators of Erythroid Enucleation. PLoS ONE *10*, e0142655.

Xu, J., Peng, C., Sankaran, V.G., Shao, Z., Esrick, E.B., Chong, B.G., Ippolito, G.C.,

Fujiwara, Y., Ebert, B.L., Tucker, P.W., et al. (2011). Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. Science *334*, 993–996.

Xu, P., Duong, D.M., Seyfried, N.T., Cheng, D., Xie, Y., Robert, J., Rush, J., Hochstrasser, M., Finley, D., and Peng, J. (2009). Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. Cell *137*, 133–145.

Yoshida, H., Kawane, K., Koike, M., Mori, Y., Uchiyama, Y., and Nagata, S. (2005). Phosphatidylserine-dependent engulfment by macrophages of nuclei from erythroid precursor cells. Nature *437*, 754–758.

Yu, D., Santos, dos, C.O., Zhao, G., Jiang, J., Amigo, J.D., Khandros, E., Dore, L.C., Yao, Y., D'Souza, J., Zhang, Z., et al. (2010). miR-451 protects against erythroid oxidant stress by repressing 14-3-3zeta. Genes Dev *24*, 1620–1633.

Zhang, J., Socolovsky, M., Gross, A.W., and Lodish, H.F. (2003). Role of Ras signaling in erythroid differentiation of mouse fetal liver cells: functional analysis by a flow cytometry-based novel culture system. Blood *102*, 3938–3946.

Zhang, L., Huang, N.-J., Chen, C., Tang, W., and Kornbluth, S. (2012). Ubiquitylation of p53 by the APC/C inhibitor Trim39. Proceedings of the National Academy of Sciences *109*, 20931–20936.

Zhang, L., Flygare, J., Wong, P., Lim, B., and Lodish, H.F. (2011). miR-191 regulates mouse erythroblast enucleation by down-regulating Riok3 and Mxi1. Genes Dev *25*, 119–124.

Zhao, B., Mei, Y., Schipma, M.J., Roth, E.W., Bleher, R., Rappoport, J.Z., Wickrema, A., Yang, J., and Ji, P. (2016). Nuclear Condensation during Mouse Erythropoiesis Requires Caspase-3-Mediated Nuclear Opening. Developmental Cell *36*, 498–510.

Zhao, G., Yu, D., and Weiss, M.J. (2010). MicroRNAs in erythropoiesis. Current Opinion in Hematology *17*, 155–162.

Zheng, N., Wang, P., Jeffrey, P.D., and Pavletich, N.P. (2000). Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. Cell *102*, 533–539.

Zhu, X., Wang, Y., Pi, W., Liu, H., Wickrema, A., and Tuan, D. (2012). NF-Y recruits both transcription activator and repressor to modulate tissue- and developmental stage-specific expression of human γ-globin gene. PLoS ONE *7*, e47175.